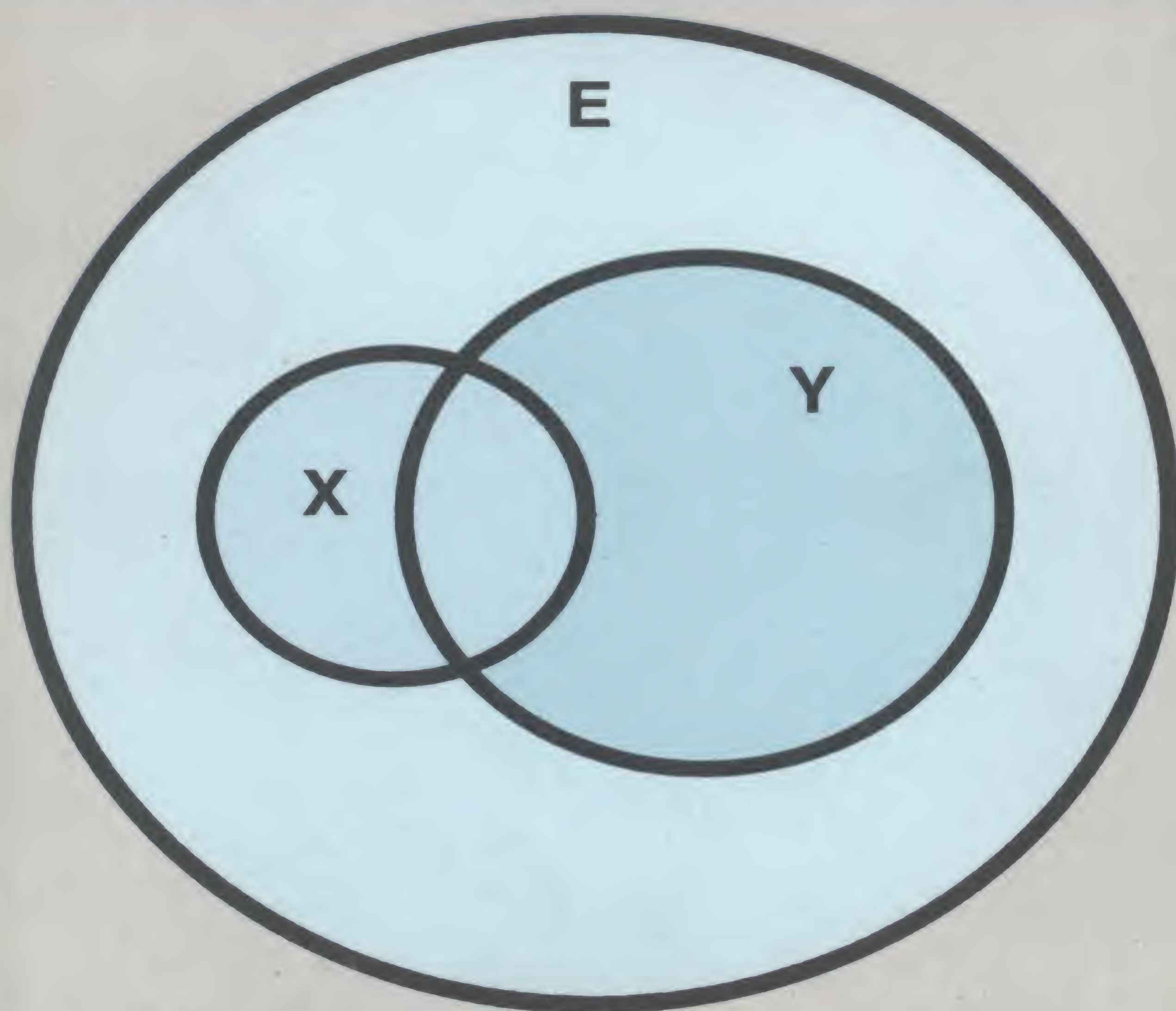
serie de matemática

monografía no. 26

ALGEBRA ELEMENTAL

Secretaría General de la
Organización de los Estados Americanos
Programa Regional de Desarrollo Científico y Tecnológico



ALGEBRA ELEMENTAL

por

Leopoldo Nachbin
Centro Brasileiro de Pesquisas Físicas
Río de Janeiro R.J., BRASIL

y

University of Rochester
Rochester, Nueva York, EE.UU.

Traducida al español por:
César E. Silva
Williams College
Williamstown, Massachusetts, EE.UU.

Secretaría General de la
Organización de los Estados Americanos
Programa Regional de Desarrollo Científico y Tecnológico
Washington, D.C. - 1986

Esta monografía ha sido preparada para su publicación en el Departamento de Asuntos Científicos y Tecnológicos de la Secretaría General de la Organización de los Estados Americanos.

Editora: Eva V. Chesneau

Asesor Técnico: César Quiroz
University of Rochester
Rochester, Nueva York, EE.UU.

A los lectores

El programa de monografías científicas es un aspecto de la vasta labor de la Organización de los Estados Americanos, a cargo del Departamento de Asuntos Científicos y Tecnológicos de la Secretaría General de dicha Organización, a cuyo financiamiento contribuye en forma importante el Programa Regional de Desarrollo Científico y Tecnológico.

Concebido por los Jefes de Estado Americanos en su Reunión celebrada en Punta del Este, Uruguay, en 1967, y cristalizado en las deliberaciones y mandatos de la Quinta Reunión del Consejo Interamericano Cultural, llevada a cabo en Maracay, Venezuela, en 1968, el Programa Regional de Desarrollo Científico y Tecnológico es la expresión de las aspiraciones preconizadas por los Jefes de Estado Americanos en el sentido de poner la ciencia y la tecnología al servicio de los pueblos latinoamericanos.

Demostrando gran visión, dichos dignatarios reconocieron que la ciencia y la tecnología están transformando la estructura económica y social de muchas naciones y que, en esta hora, por ser instrumento indispensable de progreso en América Latina, necesitan un impulso sin precedentes.

El Programa Regional de Desarrollo Científico y Tecnológico es un complemento de los esfuerzos nacionales de los países latinoamericanos y se orienta hacia la adopción de medidas que permitan el fomento de la investigación, la enseñanza y la difusión de la ciencia y la tecnología; la formación y perfeccionamiento de personal científico; el intercambio de informaciones, y la transferencia y adaptación a los países latinoamericanos del conocimiento y las tecnologías generadas en otras regiones.

En el cumplimiento de estas premisas fundamentales, el programa de monografías representa una contribución directa a la enseñanza de las ciencias en niveles educativos que abarcan importantísimos sectores de la población y, al mismo tiempo, propugna la difusión del saber científico.

La colección de monografías científicas consta de cuatro series, en español y portugués, sobre temas de física, química, biología y matemática. Desde sus comienzos, estas obras se destinaron a profesores y alumnos de ciencias de los primeros años de la universidad; de éstos se tiene testimonio de su buena acogida.

Este prefacio brinda al Programa Regional de Desarrollo Científico y Tecnológico de la Secretaría General de la Organización de los

Estados Americanos la ocasión de agradecer al doctor Leopoldo Nachbin, autor de esta monografía, y a quienes tengan el interés y buena voluntad de contribuir a su divulgación.

Julio de 1986

A la memoria de mi cuñada
Maria Regina de Meis

INDICE

Página

A los lectores v

Prólogo ix

CAPITULO 1. CONJUNTOS Y FUNCIONES 1

§ 1. Conjuntos y Elementos 1
§ 2. Subconjuntos 2
§ 3. Algebra de Conjuntos 5
§ 4. Funciones 11
§ 5. Imágenes Directas e Inversas 13
§ 6. Funciones Compuestas 17
§ 7. Relaciones de Equivalencia 24
§ 8. Espacios Cocientes 28
§ 9. Productos Cartesianos Finitos 32
§ 10. Indices 39
§ 11. Uniones e Intersecciones Arbitrarias 40
§ 12. Productos Cartesianos Arbitrarios 42

CAPITULO 2. GRUPOS 45

§ 1. Grupos Aditivos 45
§ 2. Grupos Multiplicativos 51
§ 3. Subgrupos 57
§ 4. Homomorfismos e Isomorfismos 66
§ 5. Múltiplos y Potencia 73

CAPITULO 3. ANILLOS 79

§ 1. Anillos Conmutativos 79
§ 2. Anillos Arbitrarios 83
§ 3. Subanillos 88
§ 4. Anillos con Unidad 91
§ 5. Homomorfismos e Isomorfismos 95
§ 6. Cuerpos Conmutativos 100

CAPITULO 4. ESPACIOS VECTORIALES Y ALGEBRA LINEAL 105

§ 1. Espacios Vectoriales 105
§ 2. Subespacios Vectoriales 111
§ 3. Aplicaciones Lineales e Isomorfismos 112
§ 4. Espacios Vectoriales de Dimensión Finita 115

Página

CAPITULO 5. ORDEN 123

§ 1. Orden Total 123
§ 2. Orden 126
§ 3. Reticulados 133
§ 4. Algebras de Boole 138

PROLOGO

Este texto elemental sobre los fundamentos del álgebra está dirigido a los estudiantes de primer año de todas las universidades de América Latina, no sólo en lo que concierne a matemática, sino también a campos donde el álgebra tiene aplicaciones, a saber: física, ingeniería, informática, economía, biología, estadística, etc. También está dirigido a los profesores de escuelas secundarias y universidades que quieran actualizar sus conocimientos en esta disciplina.

De modo particular quiero agradecer al matemático peruano profesor César Silva, de Williams College y de la Universidad de Rochester, por la traducción al español del manuscrito original de mi monografía, así como al matemático peruano César Quiroz, también de la Universidad de Rochester, por su colaboración en dicha traducción. Agradezco también a la Universidad de Rochester por el apoyo brindado durante la redacción de este trabajo.

1

CONJUNTOS Y FUNCIONES

Las nociones de conjunto y función desempeñan un papel fundamental en la matemática de nuestros días, incluso en algunos de sus aspectos más elementales. Ambas fueron introducidas en el siglo pasado en una forma que, en esencia, es la misma que consideramos hoy en día. Es interesante observar que fue la importante teoría de las series trigonométricas la que llevó a Cantor a estudiar sistemáticamente la teoría de los conjuntos, y a Dirichlet a ampliar el concepto de función formulado por sus antecesores y contemporáneos.

Los conjuntos y las funciones constituyen los dos elementos básicos en términos de los cuales trataremos de formular otras nociones.

§ 1. CONJUNTOS Y ELEMENTOS

Una de las nociones primitivas de la matemática es la de conjunto. Con esto queremos decir que nos limitaremos a atribuir al término *conjunto* su sentido usual de colección de objetos o elementos y no pretenderemos definirlo a partir de otros conceptos matemáticos. Por conveniencia, haremos uso también del término *colección* como sinónimo de conjunto, a fin de evitar la repetición poco elegante de este último en un mismo enunciado. El lector podrá cerciorarse de que realmente posee la noción correcta de conjunto al estudiar los varios ejemplos que se mencionan a lo largo de esta monografía.

Ejemplo 1. Los siguientes ejemplos de conjuntos se encuentran en las diversas etapas de generalización del concepto de número de la matemática elemental:

N = conjunto de los números enteros naturales $0, 1, 2, \ldots$;

Z = conjunto de los números enteros racionales $\ldots, -1, 0, 1, \ldots$;

Q = conjunto de los números racionales p/q, donde p y q son enteros racionales y $q \neq 0$;

R = conjunto de los números reales;

C = conjunto de los números complejos $a + bi$, donde a y b son números reales e $i = \sqrt{-1}$.

Se mantendrán siempre las notaciones N, Z, Q, R y C para designar a los conjuntos de números que acabamos de mencionar. En lo que sigue asumiremos que el lector está familiarizado con las propiedades de estos números.

Todo conjunto está constituido por elementos o puntos. Indicaremos que un elemento x pertenece a un conjunto X por medio de la notación de

Peano $x \in X$. La relación entre punto y conjunto definida de este modo tiene como nombre *relación de pertenencia*. Así, en el ejemplo precedente, tenemos que $2 \in \mathbf{N}$. De acuerdo con una convención general, un símbolo atravesado por una raya indica la negación de lo que sería representado por el símbolo sin la raya. En particular, para indicar que un elemento x no pertenece al conjunto X escribiremos $x \notin X$. Por ejemplo, $-2 \notin \mathbf{N}$. Si, en una determinada situación, fuera más cómodo afirmar que el conjunto X contiene al elemento x, haremos uso de la notación $X \ni x$. Esta convención, de invertir un símbolo sin alterar esencialmente su significado, es muy útil y será empleada en casos análogos, sin más comentarios.

Si bien la teoría general de los conjuntos no considera la naturaleza de los elementos que constituyen cada uno de los conjuntos, sí tiene en cuenta las relaciones posibles entre esos elementos y los conjuntos. Algunos autores tratan de describir el objetivo de la teoría de conjuntos afirmando que ella estudia las relaciones entre el todo y sus partes, y usan el nombre de *conjunto abstracto* para enfatizar la preocupación de abstraer la naturaleza de los elementos pertenecientes a los conjuntos considerados. Por nuestra parte, preferimos sólo llamar la atención del lector sobre este aspecto y no recurrir al calificativo *abstracto*.

Los varios conjuntos a los cuales se aplican con ventaja las ideas y métodos de esta teoría general pueden estar constituidos no sólo por
2 números, o por puntos en el sentido de la geometría elemental, como un cuadrado o un círculo, sino también por funciones o por otros conjuntos, etc.

Ejemplo 2. Una buena parte del álgebra elemental se dedica al estudio de las propiedades del conjunto de los polinomios. Los elementos de este conjunto son las funciones de la forma

$$a_0 + a_1 x + \ldots + a_n x^n$$

donde $a_0, a_1, \ldots, a_n$ y la variable independiente x son números reales (o complejos, conforme sea el caso).

Ejemplo 3. En geometría elemental, un plano P se concibe como un conjunto de puntos. Además de este conjunto P, también interviene en las consideraciones geométricas el conjunto R de las rectas del plano P. Este conjunto R posee la particularidad de que sus elementos son también conjuntos. Vemos así que un conjunto puede aparecer de modo natural como elemento o punto de otro conjunto: en este caso, cada recta del plano P es un elemento R de todas las rectas de P.

Desde ya se aconseja al lector a habituarse a considerar a los conjuntos como elementos o puntos de otros conjuntos, pues en construcciones importantes (como las de espacio cociente, grupo cociente, etc.) que se encontrarán más adelante, la falta de un recurso tan elemental dificultaría la comprensión de las mismas.

El signo de igualdad se utilizará siempre en el sentido usual de esta noción; esto es, escribiremos $x = y$ para indicar que los símbolos x

e y designan al mismo elemento. También se hará uso de una flecha $\Rightarrow$ para indicar implicación en el sentido de la lógica. Así, por ejemplo, si x, y y z fueran números reales, escribiremos

$$x < y \Rightarrow x + z < y + z$$

para indicar que la primera desigualdad implica la segunda. De acuerdo con la convención de inversión de símbolos, podemos también escribir

$$x < y \Leftarrow x + z < y + z$$

para indicar la implicación inversa. El empleo de la doble fecha permitirá indicar la equivalencia de dos proposiciones. Así, resumiendo las dos implicaciones anteriores, podemos escribir

$$x < y \Leftrightarrow x + z < y + z$$

A fin de amenizar el aspecto abstracto de ciertos raciocinios o definiciones es muy útil representar a los conjuntos en consideración por medio de figuras o porciones de la recta o del plano (y, en algunos casos, hasta por figuras imaginadas en el espacio tridimensional) (Fig. 1). Queda sobrentendido, entre tanto, que tales representaciones gráficas no deben interferir en las demostraciones y que tienen un propósito meramente ilustrativo.

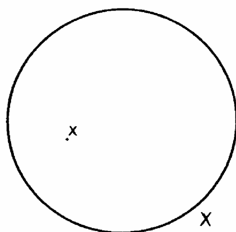


Fig. 1

Con frecuencia un conjunto aparece definido por una condición o grupo de condiciones que sus elementos deben satisfacer. Se utilizará la notación $\{x;\ c(x)\}$ para indicar al conjunto de todos los elementos x que satisfacen la condición o grupo de condiciones $c(x)$ que involucra al elemento x. Por ejemplo,

$$\{x;\ x \in \mathbf{C},\ x^5 - 1 = 0\}$$

representa al conjunto de todos los números complejos que satisfacen la ecuación $x^5 - 1 = 0$ o sea al conjunto de las raíces quintas de la unidad.

Un conjunto definido por condiciones que no pueden ser satisfechas por ningún elemento se dice *vacío* y se representa por el símbolo $\emptyset$. Por ejemplo, el conjunto siguiente

$$\{x;\ x \in \mathbf{Z},\ 2x - 1 = 0\}$$

es vacío, pues no hay ningún entero racional x que satisfaga la ecuación $2x - 1 = 0$. El papel del concepto de conjunto vacío en la teoría de conjuntos es similar al del número cero en aritmética.

Los enteros naturales 0, 1, 2,... sirven para contar los elementos de los conjuntos *finitos*, esto es, los conjuntos con un número finito de

elementos. El número 0 indica la ausencia de elementos pertenecientes al conjunto; en otras palabras, un conjunto vacío es considerado finito. Algunos autores hacen una distinción lógica entre un elemento x y el conjunto que se reduce a ese elemento, esto es el conjunto formado por apenas el elemento x, y emplean el símbolo $\{x\}$ para denotar este conjunto. Nosotros, sin embargo, adoptaremos una actitud más simplista y utilizaremos el mismo símbolo para designar tanto un elemento como al conjunto constituído por sólo ese elemento

§ 2. SUBCONJUNTOS

Dos conjuntos cualesquiera pueden ser comparados por la relación de inclusión. Diremos que un conjunto X es una parte del conjunto Y, o que X está contenido en Y o variantes similares, si todo elemento perteneciente a X también pertenece a Y. Escribiremos, entonces, $X \subset Y$ y diremos que X es un *subconjunto* de Y o que Y es un *superconjunto* de X (Fig. 2). Así, en el ejemplo 1 del párrafo precedente, tenemos que $\mathbf{N} \subset \mathbf{Z}$, $\mathbf{Z} \subset \mathbf{Q}$, $\mathbf{Q} \subset \mathbf{R}$ y $\mathbf{R} \subset \mathbf{C}$; tenemos también que $\mathbf{Z} \not\subset \mathbf{N}$. La relación $X \subset Y$ entre dos conjuntos se denomina *relación de inclusión*. Ella goza, entre otras, de las siguientes propiedades básicas:

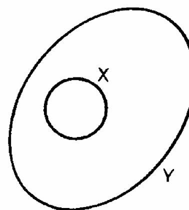


Fig. 2

4

Proposición 1. Cualesquiera que sean los conjuntos X, Y y Z, se tiene que:

1) $X \subset X$;

2) si $X \subset Y$ e $Y \subset Z$, entonces $X \subset Z$;

3) si $X \subset Y$ e $Y \subset X$, entonces $X = Y$.

Demostración. La primera propiedad es consecuencia directa de la definición de inclusión. Establezcamos la segunda propiedad. Para mostrar que $X \subset Z$, tenemos que probar que $x \in X$ implica $x \in Z$. Ahora bien, $x \in X$ implica $x \in Y$, puesto que $X \subset Y$. Además, $x \in Y$ implica $x \in Z$, pues $Y \subset Z$, como se quería demostrar. La tercera propiedad es consecuencia directa del concepto de conjunto ya presentado, de la definición de inclusión y del sentido que atribuimos a la igualdad en el párrafo anterior. QED

Las propiedades de la inclusión, que acaban de establecerse, se denominan (respectivamente) *propiedad reflexiva, ley transitiva* y *criterio de igualdad*. Nótese la evidente analogía entre esas tres propiedades y las siguientes propiedades de orden entre los números reales:

1) $x \leq x$;

2) si $x \leq y$ e $y \leq z$, entonces $x \leq z$;

3) si $x \leq y$ e $y \leq x$, entonces $x = y$

Como se verá más adelante, existen ventajas en establecer tal paralelo entre la inclusión y el orden. Entre tanto, es necesario notar que esta analogía no se puede llevar demasiado lejos. Por ejemplo, en el caso del orden entre números reales, sabemos que, dados x e y o $x \leq y$ o bien $y \leq x$. En el caso de la inclusión no es cierto que, dados dos conjuntos X e Y, tenga que ser $X \subset Y$ o $Y \subset X$. Basta tomar X como el conjunto de los enteros pares e Y como el conjunto de los enteros impares. Algunos autores expresan este hecho diciendo que la relación de inclusión es un *orden parcial*.

Cabe notar que la propiedad 1) de la proposición precedente expresa que todo conjunto es parte de sí mismo. Por eso, se dice que X es un *subconjunto propio* de Y si $X \subset Y$, pero $X \neq Y$. Se escribe, entonces, $X \underset{\neq}{\subset} Y$. Así, en el ejemplo 1 del § 1 se tiene que $\mathbf{N} \underset{\neq}{\subset} \mathbf{Z}$.

Un conjunto vacío está contenido siempre en cualquier otro conjunto. En efecto, para mostrar que una inclusión $X \subset Y$ es falsa, tenemos que (por la definición de inclusión) exhibir un elemento $x \in X$ tal que $x \in Y$. Por lo tanto, si la inclusión $\phi \subset Y$ fuera falsa debería haber un elemento $x \in \phi$ tal que $x \notin Y$, lo que es absurdo, pues ϕ es vacío. Luego $\phi \subset Y$. Tenemos aquí un ejemplo de la actitud que adoptaremos --tan corriente en matemática-- y que consiste en considerar como verdadera o satisfecha toda proposición o condición cuya negación sea absurda.

Todo conjunto X determina otro conjunto $\mathcal{P}(X)$, a saber: el conjunto de todas las partes de X. Notemos que los elementos (o puntos) de $\mathcal{P}(X)$ son las partes de X; o sea, $Y \in \mathcal{P}(X)$ es sinónimo de $Y \subset X$. Notemos también que el mismo X y su parte vacía son elementos del conjunto $\mathcal{P}(X)$. He ahí otro ejemplo de la necesidad, ya mencionada a propósito del ejemplo 3 del § 1, de considerar conjuntos cuyos elementos son también compuestos.

Ejercicio

1) Demostrar que si X es un conjunto finito con n elementos, entonces $\mathcal{P}(X)$ es finito también y tiene 2^n elementos.

§ 3. ALGEBRA DE CONJUNTOS

El estudio de las propiedades algebraicas de las operaciones de unión, intersección y complementación constituye la llamada *álgebra de conjuntos*, que posee notables analogías formales con el álgebra de ciertas operaciones usadas en la lógica de las proposiciones y estudiadas por Boole. Más tarde volveremos sobre este punto. Por el momento nos limitaremos a abordar los rudimentos del álgebra de conjuntos, indispensables para el presente capítulo.

Llámase *unión* de dos conjuntos X e Y (y se representa por $X \cup Y$) a la colección de los elementos que pertenecen por lo menos a uno de los dos conjuntos X e Y. La *intersección* de X e Y se define como la colección de los elementos que pertenecen a la vez a X y a Y, y se representa por $X \cap Y$. La *diferencia* $X - Y$ es el conjunto de todos los elementos que pertenecen a X, pero no a Y. Estas tres nociones son análogas a

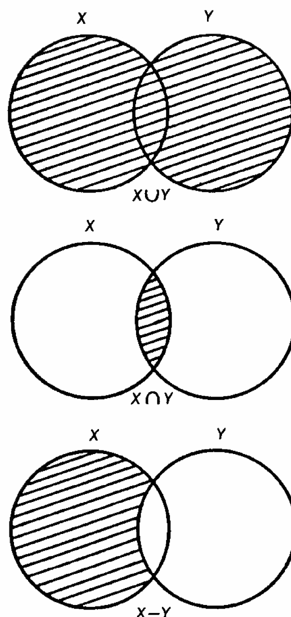


Fig. 3

la suma, producto y diferencia entre números del álgebra elemental y, por eso, algunos autores usan las notaciones $X + Y$, XY y $X - Y$ para designarlas. Estas analogías, entre tanto, no son suficientemente fuertes para imponer esta última notación. Además, existe (como veremos) una cierta dualidad entre las nociones de unión e intersección, que vamos a intentar poner de manifiesto mediante el empleo de las notaciones $\cup$ y $\cap$.

Ejemplo 1. Sean

$$X = \{x;\ x \in \mathbf{R},\ 0 \leq x < 2\},\quad Y = \{x;\ x \in \mathbf{R},\ 1 < x < 3\}$$

(según la notación introducida en el § 1). Entonces

$$X \cup Y = \{x;\ x \in \mathbf{R},\ 0 \leq x < 3\}$$

$$X \cap Y = \{x;\ x \in \mathbf{R},\ 1 < x < 2\}$$

$$X - Y = \{x;\ x \in \mathbf{R},\ 0 \leq x \leq 1\}$$

Las nociones de unión e intersección se extienden fácilmente al caso de un número finito no nulo n de conjuntos $X_1, X_2, \ldots, X_n$. Se llama *unión* de tales conjuntos al conjunto de los elementos que pertenecen a, por lo menos, uno de los conjuntos $X_1, X_2, \ldots, X_n$. Esta unión se representa por

$$X_1 \cup X_2 \ldots \cup X_n \text{ o } \cup_{i=1}^{n} X_i \text{ o } \cup_i X_i$$

Análogamente, la intersección es el conjunto de los elementos que pertenecen a todos los conjuntos X_1, X_2, ..., X_n. Las siguientes notaciones se utilizan para denotar la intersección

$$X_1 \cap X_2 \cap \ldots \cap X_n \text{ o } \cap_{i=1}^{n} X_i \text{ o } \cap_i X_i$$

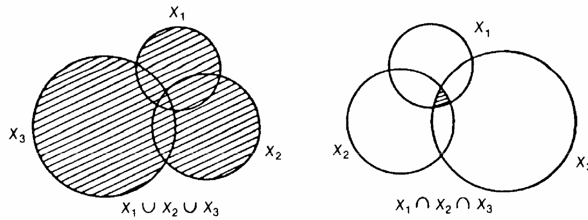


Fig. 4

En el caso particular en que $Y \subset X$, la diferencia $X - Y$ se denomina *complemento* de Y con respecto de X. Cuando el conjunto X en relación con el cual se calcula este complemento queda sobrentendido, se acostumbra representar $X - Y$ por $\complement Y$. Por ejemplo, el complemento del conjunto **Q** de los números racionales en relación con el conjunto **R** de los números reales es el conjunto de los números irracionales.

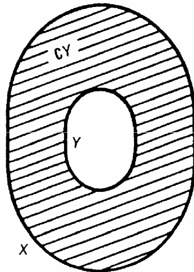


Fig. 5

Se dice que dos partes son *disjuntas* cuando no existe ningún elemento común a las mismas, o sea cuando su intersección es vacía. En el caso contrario, se dice que las dos partes se *intersecan*.

Entre las propiedades más útiles de la unión y de la intersección figuran las siguientes:

Proposición 1. Para cualesquiera conjuntos X, Y, Z, se tiene:

1) $X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$,
$X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)$,

2) $X \cup Y = Y \cup X$,
$X \cap Y = Y \cap X$,

3) $X \cup (Y \cup Z) = (X \cup Y) \cup Z$,
$X \cap (Y \cap Z) = (X \cap Y) \cap Z$,

4) $Y \subset X$ si, y sólo si, $X \cup Y = X$,
$X \subset Y$ si, y sólo si, $X \cap Y = X$.

Demostración. Establezcamos la primera parte de la propiedad 1). En virtud del criterio de igualdad (proposición 1 § 2), debemos mostrar que

$$X \cap (Y \cup Z) \subset (X \cap Y) \cup (X \cap Z) \quad \text{y} \quad (X \cap Y) \cup (X \cap Z) \subset X \cap (Y \cap Z).$$

Demostremos, pues, la primera inclusión. Si

$$t \in X \cap (Y \cup Z),$$

entonces $t \in X$ y $t \in (Y \cup Z)$. Esta última parte puede desdoblarse en dos casos: $t \in Y$ o $t \in Z$. Fijemos ideas suponiendo que $t \in Y$. Como, entonces, $t \in X$ y $t \in Y$, vemos que $t \in X \cap Y$, de donde

$$t \in (X \cap Y) \cup (X \cap Z),$$

como queríamos obtener. El otro caso, $t \in Z$, es perfectamente análogo. Establezcamos ahora la segunda inclusión. El raciocinio es exactamente el inverso del que acabamos de hacer. Si

$$t \in (X \cap Y) \cup (X \cap Z),$$

entonces $t \in X \cap Y$ o $t \in X \cap Z$. Fijemos ideas suponiendo que $t \in X \cap Y$, o sea, $t \in X$ y $t \in Y$. Ahora, $t \in Y$ implica $t \in Y \cup Z$, lo que con $t \in X$, resulta

$$t \in X \cap (Y \cup Z)$$

como queríamos. El caso en que $t \in X \cap Z$ se trata de modo semejante. Queda así establecida la primera parte de 1). En la práctica se puede dar la siguiente disposición esquemática del raciocinio hecho:

$$t \in X \cap (Y \cup Z) \Leftrightarrow t \in X,\ t \in Y \cup Z \Leftrightarrow$$

$$\Leftrightarrow \begin{Bmatrix} t \in X, & t \in Y \Leftrightarrow & t \in X \cap Y \\ t \in X, & t \in Z \Leftrightarrow & t \in X \cap Z \end{Bmatrix} \Leftrightarrow t \in (X \cap Y) \cup (X \cap Z).$$

El resto de la proposición puede establecerse de modo perfectamente análogo y por eso se deja a cargo del lector. Observemos sólo que para demostrar la primera parte de 3) puede preferirse establecer las siguientes relaciones útiles:

$$X \cup (Y \cup Z) = X \cup Y \cup Z$$

$$(X \cup Y) \cup Z = X \cup Y \cup Z$$

De igual manera, la segunda parte de 3) es consecuencia de

$$X \cap (Y \cap Z) = X \cap Y \cap Z$$

$$(X \cap Y) \cap Z = X \cap Y \cap Z.$$

QED

La propiedad expresada por 1) se llama *ley distributiva*. La primera parte es la distributividad de la intersección con relación a la

unión y la segunda es la distributividad de la unión con relación a la intersección. Recordemos a este respecto que en álgebra elemental la distributividad del producto con respecto a la suma se expresa por

$$X \cdot (Y + Z) = X \cdot Y + X \cdot Z$$

de donde se deduce la primera parte de 1) mediante la sustitución de + por $\cup$ y de $\cdot$ por $\cap$, y la segunda parte sustituyendo + por $\cap$ y $\cdot$ por $\cup$. De allí el nombre de *ley distributiva*. Las propiedades expresadas por 2) y 3) se denominan, respectivamente, *ley conmutativa* y *ley asociativa*, de nuevo por analogía con las bien conocidas conmutatividad y asociatividad de la suma y del producto en álgebra elemental. Finalmente, 4) toma el nombre de *ley de absorción*, pues ella muestra cómo, por unión o intersección, una parte puede absorber a otra.

Consideremos un conjunto universal I con relación al cual se convenga en tomar los complementos. Entre las propiedades más útiles relativas al complemento figuran las siguientes:

Proposición 2. Para cualesquiera conjuntos $X,\ Y \subset I$, se tiene:

1) $\complement(X \cup Y) = \complement X \cap \complement Y, \qquad \complement(X \cap Y) = \complement X \cup \complement Y,$

2) $X \cap \complement X = \phi, \qquad X \cup \complement X = I,$

3) $\complement(\complement X) = X.$

Demostración. La primera parte de 1) puede establecerse esquemáticamente del siguiente modo:

$$\begin{aligned} t \in \complement(X \cup Y) &\Leftrightarrow t \in I,\ t \notin X \cup Y \Leftrightarrow \\ &\Leftrightarrow t \in I,\ t \notin X,\ t \notin Y \Leftrightarrow \\ &\Leftrightarrow t \in \complement X,\ t \in \complement Y \Leftrightarrow \\ &\Leftrightarrow t \in \complement X \cap \complement Y. \end{aligned}$$

El resto de la proposición puede establecerse de modo perfectamente análogo y por eso se deja a cargo del lector. Las propiedades enumeradas en la proposición se denominan, respectivamente, *ley de dualidad, ley de complementación* y *ley de involución*. QED

Las propiedades de la unión, intersección y complemento enunciadas en las proposiciones 1 y 2 son, en cierto sentido, las propiedades fundamentales de esas operaciones, o sea que no fueron escogidas meramente al acaso. La formulación exacta del sentido en el cual las referidas propiedades son realmente fundamentales requeriría nociones de teoría de reticulados y de álgebras de Boole que, por motivos metodológicos, preferimos discutir más adelante.

No es necesario poseer un agudo espíritu de observación para advertir, al examinar los enunciados de las proposiciones 1 y 2, la presencia de una *dualidad*. Tal dualidad salta a la vista tan fácilmente que

nadie vacilaría en formularla en los siguientes términos generales: "De toda igualdad válida entre conjuntos arbitrarios se deduce otra fórmula también válida, siempre que se permuten los signos $\cup$ y $\cap$, se permuten los conjuntos ϕ y I y se conserve el signo **C**". No sería tampoco difícil prever que "de toda inclusión válida entre conjuntos arbitrarios, se deduce otra inclusión también válida, con tal que se permuten los dos miembros y se proceda como en el caso de las igualdades". No es nuestra intención precisar los conceptos de igualdad e inclusión válidos entre conjuntos arbitrarios para luego demostrar los dos principios de dualidad que acaban de ser enunciados --como sería indispensable hacerlo desde el punto de vista del rigor estricto-- pues los mismos son lo suficientemente claros para que no dejen duda en cuanto a su aplicación correcta.

Ejercicios

1) Establecer las siguientes propiedades:

$X \cup \phi = X$, $\qquad X \cap \phi = \phi$,

$X \cup X = X$, $\qquad X \cap X = X$,

$X \subset X \cup Y$, $\qquad X \cap Y \subset X$,

$X \subset Z,\ Y \subset Z \Leftrightarrow X \cup Y \subset Z$, $\qquad Z \subset X,\ Z \subset Y \Leftrightarrow Z \subset X \cap Y$,

$X \subset Y \Rightarrow X \cup Z \subset Y \cup Z$, $\qquad X \subset Y \Rightarrow X \cap Z \subset Y \cap Z$,

$X \cap (X \cup Y) = X$, $\qquad X \cup (X \cap Y) = X$.

2) Establecer la siguiente *ley de corte:*

$X \cup Z = Y \cup Z$ y $X \cap Z = Y \cap Z$ equivalen a $X = Y$.

3) Establecer las siguientes propiedades, en las cuales el complemento se toma con respecto a un conjunto I.

$\mathbf{C}\phi = I$ $\qquad \mathbf{C}I = \phi$,

$X \cap Y = \phi \Leftrightarrow X \subset \mathbf{C}Y \Leftrightarrow Y \subset \mathbf{C}X$,

$X \cup Y = I \Leftrightarrow \mathbf{C}Y \subset X \Leftrightarrow \mathbf{C}X \subset Y$.

4) Demostrar la siguiente identidad:

$(X \cap Y) \cup (X \cap \mathbf{C}Y) \cup (\mathbf{C}X \cap Y) \cup (\mathbf{C}X \cap \mathbf{C}Y) = I$

demostrando también su significado en un gráfico de los conjuntos X, Y e I. Escribir la identidad dual.

5) Mostrar que la diferencia se expresa mediante la intersección y la complementación por $X - Y = X \cap \mathbf{C}Y$.

§ 4. FUNCIONES

La noción de función --formulada por Dirichlet para satisfacer los requerimientos de la teoría de las series trigonométricas-- es la de una correspondencia que a cada número real le asocia otro número real. En matemática, entre tanto, ocurren otros tipos de correspondencia, en las cuales, por ejemplo, se asocian números reales a funciones o funciones a otras funciones, etc. La noción general de función, que así se torna indispensable introducir, difiere de la noción debida a Dirichlet, en que las correspondencias consideradas no se limitan a asociar números reales a números reales, sino que pueden actuar entre los elementos de dos conjuntos de naturaleza cualquiera.

Una función definida en el conjunto X con valores en el conjunto Y es una correspondencia que a cada punto de X le asocia un punto de Y. También se usan como sinónimos de "función de X en Y" los términos *aplicación* o *transformación* de X en Y. Si se designa una función mediante una letra, f por ejemplo, se representará por $f(x)$, o simplemente fx, el valor de f en el punto $x \in X$; o sea, el punto de Y que corresponde a x por f. El conjunto X recibe el nombre de *dominio* de la función e Y es su *contradominio*. No se excluye el caso en que los conjuntos X e Y coincidan: en este caso, f se llama una aplicación de X en sí mismo. Si se desea indicar el dominio o contradominio de la función, se escribe $f : X \to Y$ o $X \xrightarrow{f} Y$. Asimismo, es muy cómodo indicar una función mediante el símbolo $x \mapsto f(x)$ mostrando el elemento de $f(x)$ que corresponde a x. Observemos que la noción de función $f : X \to Y$ abarca dos conjuntos y una correspondencia. Por tanto, de acuerdo con el sentido atribuido en el § 1 a la noción de igualdad, dos funciones $f : X \to Y$ y $f' : X' \to Y'$ son iguales si, y sólo si, $X = X'$, $Y = Y'$ y $f(x) = f'(x)$ cualquiera que sea $x \in X$.

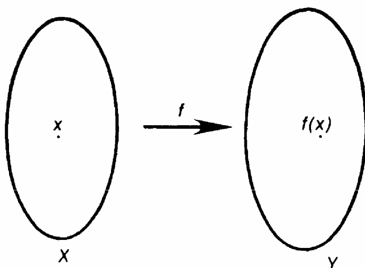


Fig. 6

Ejemplo 1. La correspondencia dada por $x \mapsto x^2$, o sea la que asocia a cada $x \in \mathbf{R}$ su cuadrado $x^2 \in \mathbf{R}$, define una función con valores en $\mathbf{R}$, donde $\mathbf{R}$ designa el conjunto de los números reales. De una manera general, toda función con valores en $\mathbf{R}$, definida en $\mathbf{R}$ (o sólo en un subconjunto de $\mathbf{R}$, como un intervalo), se denomina *función real de variable real*. Este es el primer tipo de función que se encuentra en matemática elemental. La colección de todas las funciones reales de variable real constituye un ejemplo importante de conjunto.

Ejemplo 2. Indiquemos con C el conjunto de las funciones reales continuas en un intervalo cerrado acotado $[a, b]$ de $\mathbf{R}$. La correspondencia que a cada función $f \in C$ le asocia su integral $\int_a^b f(x)\, dx$ constituye un ejemplo de función definida en C con valores en $\mathbf{R}$. Para evitar

el uso de expresiones como "la integral es una función de la función integrada", que podría parecer confusa, se suele emplear el término *funcional* en vez de función para designar correspondencias que asocian números a funciones. Así, la integral es una funcional de la función integrada.

Ejemplo 3. Consideremos el conjunto D de las funciones reales de variable real, derivables en **R**, y el conjunto F de las funciones reales de variable real. La correspondencia $f \mapsto f'$ que a toda $f \in D$ asocia su derivada f' constituye un ejemplo de función definida en D, con valores en F. En este ejemplo y en algunas situaciones similares, se prefiere usar el término *operador* en vez de función para designar la correspondencia considerada.

Se llama la atención del lector al hecho de que se usará el término *función* para lo que, en la terminología de otros autores, se denomina *función unívoca*. Con esto no se excluyen las llamadas *funciones plurívocas*, ya que, en realidad, éstas se reducen a aquéllas. En efecto, sea f una función plurívoca de un conjunto X en un conjunto Y, esto es, una correspondencia que a cada $x \in X$ asocia varios puntos $y \in Y$. Para cada punto $x \in X$ representaremos por $f(x)$ al conjunto de los puntos $y \in Y$ que corresponden a x mediante f. Entonces $f(x)$ es un subconjunto de Y y la función plurívoca f puede interpretarse como una función unívoca que a cada elemento $x \in X$ hace corresponder una parte $f(x) \subset Y$, o sea un elemento $f(x) \in \mathcal{P}(Y)$, donde $\mathcal{P}(Y)$ designa el conjunto de las partes de Y (véase el § 2). En otros términos, una función plurívoca de X en Y podrá ser definida como una función de X en $\mathcal{P}(Y)$.

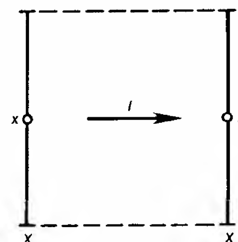


Fig. 7

Se dijo ya que todo conjunto X da lugar a otro conjunto, a saber, a la colección $\mathcal{P}(X)$ de todas las partes de X. En el caso de funciones, cabe una observación parecida. Si X e Y designan dos conjuntos, puede ser útil en una situación determinada considerar el conjunto de todas las funciones definidas en X con valores en Y: este conjunto total de funciones se suele representar por el símbolo Y^X (el origen de esta notación se indica en el subsiguiente ejercicio 1).

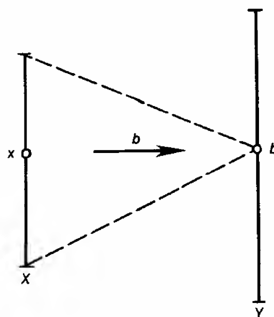


Fig. 8

Se da el nombre de *transformación identidad* de un conjunto X y se representa por la letra I, a

la transformación de X en sí misma, que a cada punto $x \in X$ asocia el mismo x. Por tanto, $I : X \to X$ es tal que $I(x) = x$, para todo $x \in X$. Así, en el caso del conjunto **R** de los números reales, la transformación identidad de **R** es una función cuyo gráfico en el plano es la bisectriz $y = x$.

Si X e Y designan dos conjuntos, toda función de X en Y que a todos los puntos de X asocia un mismo punto de Y se dice constante. En otras palabras, si b fuese un punto de Y, la función de X en Y que a cada punto $x \in X$ asocia el punto b recibe el nombre de *función constante* y se representa por el mismo símbolo b. Luego, $b : X \to Y$ es tal que $b(x) = b$ para todo $x \in X$.

Ejercicio

1) Las funciones de un conjunto finito X de m elementos en otro conjunto finito Y de n elementos constituyen un conjunto finito de n^m elementos. (De aquí la notación ya indicada, Y^X.)

§ 5. IMAGENES DIRECTAS E INVERSAS

Se acostumbra también dar el nombre de *conjunto de definición* al dominio X de una función $f : X \to Y$. Se llama conjunto de valores de f a

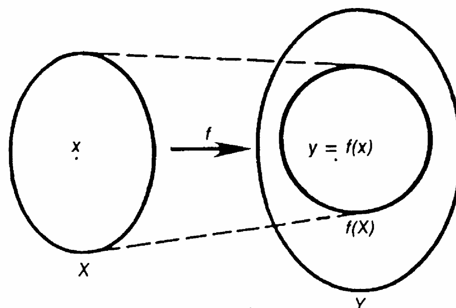


Fig. 9

la colección de puntos $y \in Y$ para los cuales existe por lo menos un punto $x \in X$ tal que $f(x) = y$; o como también se acostumbra decir, al conjunto de los puntos de la forma $f(x)$, $x \in X$. El conjunto de los valores es una parte o subconjunto del contradominio Y y puede coincidir o no con este contradominio. Así, en el ejemplo 1 del párrafo precedente, la función $x \mapsto x^2$ de **R** en **R** tiene como conjunto de sus valores al conjunto **R**+ de los números reales no negativos, que es una parte propia del contradominio **R**. Por otro lado, el conjunto de los valores de la función definida por la correspondencia descrita en el ejemplo 2 coincide con el contradominio **R**: en efecto, todo número real k es igual a la integral de, por lo menos, una función real continua en $[a, b]$; basta tomar la función constante $k / (b - a)$, siempre que $a \neq b$.

El hecho de que el conjunto de los valores de una función sea una parte propia de su contradominio se traduce en que, en cierto sentido,

el contradominio es innecesariamente amplio para el tipo de correspondencia que define la función. Así, en el caso del ejemplo 1 anterior, es perfectamente claro *a priori* que basta considerar la correspondencia $x \mapsto x^2$ entre **R** y **R**+ en vez de entre **R** y **R**. En el caso del ejemplo 3 del párrafo anterior, un teorema de Darboux permite mostrar que el conjunto de valores de la operación de derivación es una parte propia de F, más no es tan fácil caracterizar *a priori* tal conjunto de valores. En los cursos de cálculo infinitesimal se acostumbra referirse a las funciones que pertenecen al conjunto de valores de la operación de derivación como funciones que "poseen primitiva".

Se dice que una función $f : X \to Y$ es *sobre* Y *o suryectiva* o que f es una *suryección* cuando su conjunto de valores coincide con el contradominio Y; o sea, cuando para todo $y \in Y$ existe, por lo menos, un $x \in X$ tal que $f(x) = y$.

Si $A \subset X$ y $f : X \to Y$ es una función, se denomina *imagen directa* de A por f al conjunto de los valores que f toma en A, o sea, al conjunto de los puntos $y \in Y$ para los cuales existe, por lo menos, un $x \in A$ tal que $y = f(x)$. Esta imagen directa se representa por $f(A)$. Notemos que la

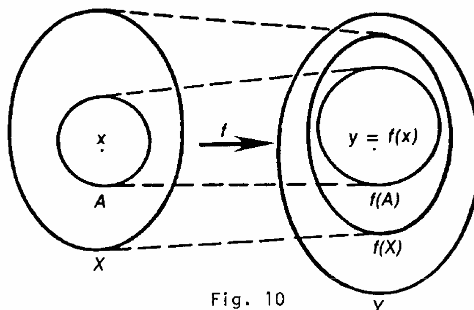


Fig. 10

imagen directa $f(X)$ de X coincide con el conjunto de los valores de f definido anteriormente. Por lo tanto, la condición de que $f : X \to Y$ sea sobre Y puede expresarse por la igualdad $f(X) = Y$. Observemos asimismo que $f(A) \subset Y$ y también que $f(A) \subset f(X)$. La imagen directa de la parte de X reducida a un punto $x \in X$ es la parte de Y reducida al punto $f(x)$; de ahí que al punto $f(x)$ también se le llama imagen directa de x.

Si $B \subset Y$, se da el nombre de *imagen inversa* de B por f al conjunto de puntos $x \in X$ cuyas imágenes directas $f(x)$ pertenecen a B. Esta imagen inversa se representa por $f^{-1}(B)$. Notemos que siempre $f^{-1}(Y) = X$ dado que todo punto $x \in X$ tiene su imagen directa $f(x)$ en Y. Observemos también que $f^{-1}(B) \subset X$.

Una diferencia relacionada con el caso de las imágenes directas es la siguiente. Si se considera una parte de Y reducida a un punto y, su imagen inversa $f^{-1}(y)$, o sea el conjunto de los puntos $x \in X$ tales que $y = f(x)$, puede ser vacía o no y, en este último caso, puede consistir de uno o más elementos. Para que $f^{-1}(y)$ no sea vacía es necesario y suficiente que y pertenezca al conjunto de los valores $f(X)$; en efecto, la

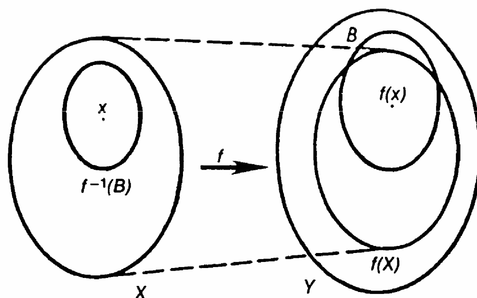


Fig. 11

existencia de por lo menos un $x \in f^{-1}(y)$ equivale a la existencia de por lo menos un x tal que $y = f(x)$. A título de aclaración, se mencionan los siguientes ejemplos. En el caso de la función $x \mapsto x^2$ de **R** en **R**, la imagen inversa de cada $y \in \mathbf{R}$ puede dar lugar a los siguientes casos: 1) si $y > 0$, esta imagen inversa consiste de dos puntos $+\sqrt{y}$, $-\sqrt{y}$; 2) si $y = 0$, la imagen inversa consiste sólo del número 0; 3) si $y < 0$ su imagen inversa es vacía. En el caso del ejemplo 3 del párrafo precedente, si f designa una función que posee primitiva, esto es, si f pertenece al conjunto de los valores de la operación de derivación, cada función perteneciente a la imagen inversa de f por esta operación, o sea cada función cuya derivada es igual a f se denomina *primitiva* de f y se representa por $\int f(x)dx$. Como es sabido, la imagen inversa de f consiste de todas las funciones de la forma $\int f(x)dx + c$, esto es, toda primitiva de f es la suma de una primitiva particular de f con una constante arbitraria.

Así como el estudio de las imágenes directas conduce a la noción de función suryectiva, el estudio de las imágenes inversas lleva a la noción de función biunívoca, que pasaremos a definir. Existe una cierta *dualidad* entre la noción de función sobre y la de función biunívoca, pero no nos detendremos aquí para formularla explícitamente; el lector atento podrá ciertamente notarla (especialmente en los enunciados de los ejercicios subsiguientes).

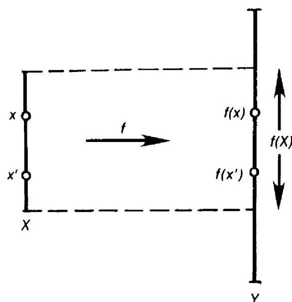


Fig. 12

Se dice que una función $f : X \to Y$ es *biunívoca* o *inyectiva*, o que f es una *inyección* de X en Y si para todo par de puntos distintos de X sus imágenes directas son distintas. En términos de la noción de imagen inversa, la inyectividad de $f : X \to Y$ se expresa diciendo que para todo $y \in Y$, la imagen inversa $f^{-1}(y)$ contiene a lo más un punto (esto es, o consiste de un punto o es vacía). En efecto, si f fuera inyectiva y x, $x' \in f^{-1}(y)$, entonces $f(x) = y$, $f(x') = y$, de donde $f(x) = f(x')$, lo que implica que $x = x'$, o sea que $f^{-1}(y)$ no puede contener dos

puntos distintos. Recíprocamente, si cada $f^{-1}(y)$ contuviera a lo máximo un punto y tuviéramos $x \neq x'$, entonces debiéramos tener $f(x) \neq f(x')$, puesto que, en caso contrario, con $y = f(x) = f(x')$, tendríamos $x, x' \in \in f^{-1}(y)$, lo que exigiría que $x = x'$.

Adviértase que la noción de función biunívoca o inyectiva como fue definida aquí no exige que la función sea sobre su contradominio, a diferencia de lo que acostumbran algunos autores. De acuerdo, entonces, con la terminología que preferimos seguir, diremos *función biunívoca sobre* cada vez que tal fuera el caso. Una función biunívoca de X sobre Y se denominará también *correspondencia biunívoca* entre X e Y. Resumiendo, por razones de claridad, tenemos, por una parte, el caso general de las aplicaciones biunívocas (o inyectivas) de X en Y y, por la otra, el caso particular de las aplicaciones biunívocas de X sobre Y o, equivalentemente, las correspondencias biunívocas entre X e Y.

Asociada a toda función biunívoca $f : X \to Y$ de X sobre Y existe su función inversa que pasaremos a definir. Dada una función $f : X \to Y$, hemos visto ya que $f^{-1}(y)$ contiene, por lo menos, un punto para todo $y \in Y$ si, y sólo si, f es sobre Y; y, además de esto, $f^{-1}(y)$ contiene a lo máximo un punto para todo $y \in Y$ si, y sólo si, f es biunívoca (o inyectiva). Combinando estas dos observaciones vemos que $f^{-1}(y)$ se reduce a un punto para todo $y \in Y$ si, y sólo si, f es biunívoca sobre Y. De acuerdo con una convención general (§ 1), se representará por $f^{-1}(y)$ al punto al que el conjunto $f^{-1}(y)$ se reduce; la función $f^{-1} : Y \to X$, que a cada punto $y \in Y$ asocia el punto $f^{-1}(y) \in X$ se denomina *función inversa* de la función dada. Esta notación sólo tiene sentido en el caso de una función biunívoca sobre. La función inversa $f^{-1}(y) : Y \to X$ es biunívoca de Y sobre X. En efecto, si $y, y' \in Y$ e $y \neq y'$, entonces $f^{-1}(y) \neq f^{-1}(y')$, ya que si escribimos $f^{-1}(y) = x$ y $f^{-1}(y') = x'$ tendremos $y = f(x)$ e $y' = f(x')$ y, entonces, $x = x'$ implicaría $y = y'$, luego f^{-1} es biunívoca (o inyectiva) en X. Además, dado cualquier $x \in X$, poniendo $y = f(x)$ se obtiene un punto $y \in Y$ tal que $x \in f^{-1}(y)$, o sea $x = = f^{-1}(y)$, luego f^{-1} es sobre X. Como toda función biunívoca de su dominio sobre su contradominio, $f^{-1} : Y \to X$ posee una función inversa $(f^{-1})^{-1} : : X \to Y$, y no presenta dificultad alguna verificar que ésta coincide con la propia función $f : X \to Y$. Las nociones de función biunívoca y función inversa son suficientemente conocidas en matemática elemental; por ejemplo, la función exponencial $x \mapsto e^x$ definida de $\mathbf{R}$ sobre el conjunto de los números reales > 0 y la función logaritmo $x \mapsto \log x$ definida de este último conjunto sobre $\mathbf{R}$ son ambas biunívocas sobre y una la inversa de la otra.

Se da el nombre de *permutación* de un conjunto X a toda aplicación biunívoca de X sobre sí mismo. A toda permutación corresponde una permutación inversa. Una *involución* es una permutación que coincide con su propia inversa. Por ejemplo, la aplicación de $\mathbf{R}$ en $\mathbf{R}$ definida por $x \mapsto ax + b$, donde $a \neq 0$, es una permutación de $\mathbf{R}$. Su inversa es la aplicación $x \mapsto (x - b)/a$. No presenta dificultad alguna verificar que la aplicación dada es una involución sólo en los dos casos siguientes: $a = -1$ o $a = 1$, $b = 0$. Otro ejemplo: si I fuese un conjunto cualquiera, la aplicación $X \mapsto \complement X$ que a toda parte $X \subset I$ asocia su complemento en I constituye una permutación de $\mathcal{P}(I)$ que es una involución.

El estudio de las imágenes directas e inversas hecho en este párrafo se repetirá más adelante para el caso de los homomorfismos entre grupos, anillos, etc., donde se podrá obtener importante información adicional dada la mayor riqueza de las estructuras.

Ejercicios

1) Una función continua $f : \mathbf{R} \to \mathbf{R}$ es sobre $\mathbf{R}$ si, y sólo si, f es ilimitada (no acotada) tanto inferiormente como superiormente. Si f fuese el cociente de dos polinomios, esta condición se verifica si, y sólo si, el grado del numerador menos el del denominador es un número entero impar positivo.

2) Sea $f : X \to Y$ una función. Probar que

$$A \subset A' \Rightarrow f(A) \subset f(A'), \qquad B \subset B' \Rightarrow f^{-1}(B) \subset f^{-1}(B'),$$

$$f(A \cup A') = f(A) \cup f(A'), \qquad f^{-1}(B \cup B') = f^{-1}(B) \cup f^{-1}(B').$$

3) Sea $f : X \to Y$ una función. Se tiene

$$f^{-1}(B \cap B') = f^{-1}(B) \cap f^{-1}(B').$$

En el caso de las imágenes directas sólo se cumple una inclusión:

$$f(A \cap A') \subset f(A) \cap f(A').$$

Para que la igualdad

$$f(A \cap A') = f(A) \cap f(A')$$

sea válida cualesquiera que sean A y A' es necesario y suficiente que f sea inyectiva en X.

4) Sea $f : X \to Y$ una función. Se tiene

$$f^{-1}(\mathbf{C}B) = \mathbf{C}\, f^{-1}(B)$$

Para que

$$f(\mathbf{C}A) = \mathbf{C}f(A)$$

sea válida para todo A es necesario y suficiente que f sea biunívoca de X sobre Y.

5) Sea $f : X \to Y$ una función. Se tiene $f(\phi) = \phi$ y, más precisamente, $f(A) = \phi$ si, y sólo si, $A = \phi$. Se tiene también $f^{-1}(\phi) = \phi$ y, más precisamente, $f^{-1}(B) = \phi$ si, y sólo si, $B \cap f(X) = \phi$. Además de esto, $f^{-1}(B) = f^{-1}\{B \cap f(X)\}$.

§ 6. FUNCIONES COMPUESTAS

Uno de los modos de combinar funciones, ya considerado en matemática elemental, consiste en aplicar una función después de la otra;

de esta manera se introducen los conceptos de *función de función* o de *función compuesta,* según que las funciones consideradas sean de una o más variables reales o complejas. Como se verá más adelante en la sección sobre los productos cartesianos, las funciones de varias variables pueden concebirse como funciones de una variable, y tal distinción entre los conceptos de función de función y de función compuesta, introducida por motivos didácticos, pasará a ser innecesaria.

Por otra parte, en la teoría elemental de las permutaciones se define un *producto* de dos permutaciones del mismo grupo de letras efectuando una permutación después de la otra. Finalmente, en varias situaciones geométricas es común estudiar el efecto del producto de dos transformaciones (o sea, la aplicación consecutiva de las mismas) sobre las figuras a fin de probar hechos, como: "el producto de dos transformaciones proyectivas es también una transformación proyectiva".

Algunos aspectos comunes a esos casos pueden ser condensados en el estudio de la noción general de producto de dos funciones. Consideremos, en efecto, dos funciones $f : X \to Y$ y $g : Y \to Z$, tales que el contradominio de la primera coincide con el dominio de la segunda. Por la primera función, a todo elemento $x \in X$ le corresponde el elemento $f(x) \in Y$. Además, a todo elemento $y \in Y$ le corresponde por la segunda función, el elemento $g(y) \in Z$; en particular, al elemento $y = f(x) \in Y$ le corresponderá, por la segunda función, el elemento $g(y) = g(f(x)) \in Z$. Haciendo corresponder directamente a todo elemento $x \in X$ el elemento $g\{f(x)\}$, definimos una función de X en Z que se representará por $gf : X \to Z$. Esta función es tal que

$$(gf)(x) = g\{f(x)\} \text{ para } x \in X.$$

Obsérvese el aspecto de ley asociativa de esta ecuación. Se dice que $gf : X \to Z$ es una *función de funciones* o, de preferencia, que gf es una función *compuesta o producto* de $f : X \to Y$ y $g : Y \to Z$ en este orden, cuando $gf : X \to Z$ ha sido obtenida a partir de $f : X \to Y$ y $g : Y \to Z$ por el proceso descrito antes.

En el caso particular en que $X = Z$ y, por tanto, $f : X \to Y$ y $g : Y \to X$ actúan en direcciones opuestas, tiene sentido no sólo formar la función compuesta $gf : X \to X$, dado que el contradominio de f y el dominio de g coinciden, sino que también puede formarse la función compuesta $fg : Y \to Y$ definida por

$$(fg)(y) = f\{g(y)\} \text{ para } y \in Y$$

pues el contradominio de g y el dominio de f también coinciden. En virtud del sentido atribuido a la igualdad de funciones (§ 4), es conveniente averiguar si $gf : X \to X$ y fg $Y \to Y$ son iguales en el caso en que $X = Y$, o sea en que $f : X \to X$ y $g : X \to X$ son ambas transformaciones del conjunto X en sí mismo. Es necesario, por tanto, notar que $gf : X \to X$ y $fg : X \to X$ pueden no ser iguales; esto es, el orden en que se efectúe el producto puede afectar el resultado. Cuando $gf = fg$, o sea cuando

$$g\{f(x)\} = f\{g(x)\} \text{ para todo } x \in X,$$

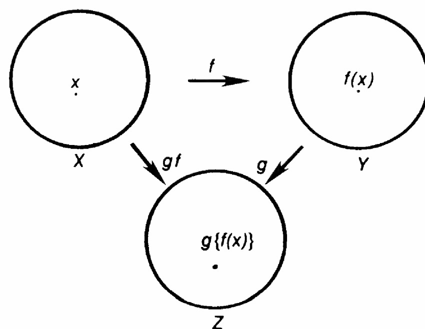


Fig. 13

se dice que f y g *conmutan*. Es éste un primer y rápido contacto con los llamados sistemas *no conmutativos*, cuya consideración por razones de simplicidad se descuida en álgebra elemental, pero que constituyen un tema central de estudio en álgebra moderna.

Ejemplo 1. Consideremos las funciones de **R** en **R** definidas por $f : x \mapsto x^2$ y $g : x \mapsto e^x$. Las funciones de **R** en **R** obtenidas por la composición de éstas en los dos órdenes posibles son $gf : x \mapsto e^{x^2}$ y $fg : x \mapsto e^{2x}$; luego, las funciones dadas no conmutan. En realidad, no es necesario considerar las funciones exponenciales, etc., para dar un ejemplo de no conmutatividad: la condición para que dos funciones lineales $x \mapsto ax + p$ y $x \mapsto bx + q$, de **R** en **R**, conmuten es que $(a - 1)q = (b - 1)p$, y basta escoger cuatro coeficientes que no cumplan esta igualdad para obtener el ejemplo deseado.

Ejemplo 2. Consideremos en un plano P, un punto fijo c. Dado un ángulo a, $-\infty < a < +\infty$, consideremos la rotación c_a del plano P en torno de c con ángulo a, o sea la correspondencia que asocia a todo punto $x \in P$ el punto $y = c_a(x) \in P$ tal que $\text{ang}(\vec{cx}, \vec{cy}) = a$ y $\text{dist}(c, x) = \text{dist}(c, y)$ si $x \neq c$, y al punto c se asocia el propio c. Cada c_a es una función de P en P. El producto de dos tales rotaciones c_a y c_b, en este orden, es la rotación c_{a+b}, o sea $c_b c_a = c_{a+b}$, lo que en mayor detalle significa que $c_b\{c_a(x)\} = c_{a+b}(x)$ para $x \in P$. Permutando los papeles de a y b se tiene asimismo que $c_a c_b = c_{a+b}$, lo que implica que $c_a c_b = c_b c_a$,

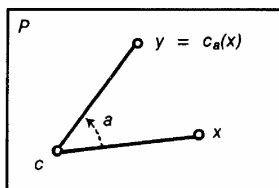


Fig. 14

o sea que dos rotaciones del plano en torno a un mismo punto siempre conmutan.

Ejemplo 3. Consideremos el conjunto $\mathcal{D}$ de las funciones reales de variable real infinitamente diferenciables en $\mathbf{R}$. A cada función $f \in \mathcal{D}$ asociemos la función $g \in \mathcal{D}$ definida por $g(x) = \int_0^x f(x)\, dx$. Representemos por $i : \mathcal{D} \to \mathcal{D}$ la correspondencia así definida: $i(f) = g$. Análogamente, a cada función $g \in \mathcal{D}$ asociemos la función $f \in \mathcal{D}$ definida por $f(x) = dg(x)/dx$. Representemos por $d : \mathcal{D} \to \mathcal{D}$ la correspondencia así definida: $d(g) = f$. Entonces $di : \mathcal{D} \to \mathcal{D}$ es la transformación identidad de $\mathcal{D}$, mientras que $id : \mathcal{D} \to \mathcal{D}$ es "casi" la transformación identidad de $\mathcal{D}$, pues a cada $g \in \mathcal{D}$ corresponde la función $i\{d(g)\} = f$ tal que $f(x) = g(x) - g(0)$.

A continuación vamos a establecer algunas propiedades bastante sencillas de la noción de producto de funciones. Conviene observar que no se excluirá el caso en que se tenga igualdad de los conjuntos que constituyen los dominios y contradominios de las funciones mencionadas, como será realmente lo que sucede cuando se considere el grupo de las permutaciones de un conjunto.

Proposición 1. Si $f : W \to X$, $g : X \to Y$ y $h : Y \to Z$ son tres funciones, entonces las funciones $h(gf)$ y $(hg)f$ de W en Z son iguales.

Demostración. Comencemos notando que los productos $gf : W \to Y$ y $hg : X \to Z$ tienen sentido, puesto que el contradominio de f es igual al dominio de g y el contradominio de g es igual al dominio de h. Por motivos similares, los productos $h(gf) : W \to Z$ y $(hg)f : W \to Z$ también tienen sentido. Estos últimos productos poseen el mismo dominio y contradominio, y para establecer su igualdad basta demostrar que

$$\{h(gf)\}(w) = \{(hg)f\}(w) \quad \text{para todo } w \in W.$$

Ahora,

$$\{h(gf)\}(w) = h\{(gf)(w)\} = h\{g[f(w)]\}$$

$$\{(hg)f\}(w) = (hg)\{f(w)\} = h[g\{f(w)\}],$$

lo que establece la igualdad deseada. QED

La proposición anterior expresa la llamada *ley asociativa* del producto de funciones. Por ella es posible definir sin ambigüedad el producto $hgf : W \to Z$ como la función $h(gf) : W \to Z$ o la función $(hg)f : W \to Z$. La extensión al caso del producto de un número finito de funciones es inmediata.

Proposición 2. Si $f : X \to Y$ es una función e $I : X \to X$ es la transformación identidad de X, entonces $fI = f$, esto es, las funciones $f : X \to Y$ y $fI : X \to Y$ son iguales. Análogamente, si $I : Y \to Y$ es la transformación identidad de Y, entonces $If = f$.

Demostración. Como el contradominio de $I : X \to X$ y el dominio de $f : X \to Y$ son iguales, tiene sentido el producto $fI : X \to Y$. Ahora bien

$$(fI)(x) = f\{I(x)\} = f(x)$$

para todo $x \in X$, lo que prueba la igualdad $fI = f$. De forma análoga se obtiene la igualdad $If = f$. QED

La proposición que acaba de demostrarse expresa, simplemente, que las transformaciones identidad desempeñan el papel de *unidad* en relación con el producto de funciones. Conviene entretanto resaltar que, con el mismo símbolo I, se indican las varias transformaciones identidad; en particular, en las ecuaciones $fI = f$ e $If = f$ del enunciado de la proposición tenemos en realidad dos transformaciones distintas, a menos que $X = Y$.

Proposición 3. Si $f : X \to Y$ es una función biunívoca de X sobre Y y $f^{-1} : Y \to X$ es su función inversa, entonces $f^{-1}f = I$ y $ff^{-1} = I$, esto es, $f^{-1}f : X \to X$ y $ff^{-1} : Y \to Y$ son iguales a $I : X \to X$ e $I : Y \to Y$, respectivamente.

Recíprocamente, si $f : X \to Y$ y $g : Y \to X$ son dos funciones tales que $gf : X \to X$ y $fg : Y \to Y$ son las transformaciones identidad de X e Y, respectivamente, entonces $f : X \to Y$ es una función biunívoca de X sobre Y, cuya inversa es igual a $g : Y \to X$.

Demostración. Comencemos estableciendo la primera parte del enunciado. Es claro que tiene sentido efectuar el producto $f^{-1}f : X \to X$. Tenemos

$$(f^{-1}f)(x) = f^{-1}\{f(x)\}$$

para todo $x \in X$, por la definición de producto. Ahora, escribiendo $f(x) = y$, vemos que, por la definición de función inversa, $f^{-1}(y) = x$. Luego,

$$(f^{-1}f)(x) = x = I(x),$$

lo que prueba la igualdad $f^{-1}f = I$. La demostración de la igualdad $ff^{-1} = I$ es análoga.

Establezcamos ahora la segunda parte del enunciado. Es claro que los productos $gf : X \to X$ y $fg : Y \to Y$ tienen sentido. Comencemos demostrando que la función f es inyectiva en X. En efecto, consideremos dos puntos cualesquiera x, $x' \in X$, tales que $x \neq x'$. Como, por hipótesis, $gf = I$, tenemos

$$g\{f(x)\} = x \quad \text{y} \quad g\{f(x')\} = x'.$$

Esto implica $f(x) \neq f(x')$, dado que $f(x) = f(x')$ implicaría $x = x'$, lo que está en contra de nuestra hipótesis. Luego, f es inyectiva en X. Pasemos ahora a mostrar que la función f es sobre Y. En efecto, consideremos un punto cualquiera $y \in Y$. Como, por hipótesis, $fg = I$, tenemos

$$f\{g(y)\} = y;$$

escribiendo, entonces, $x = g(y)$, obtenemos un punto $x \in X$ tal que $f(x) = y$ y, por tanto, f es sobre Y. Este mismo razonamiento muestra que

$g = f^{-1}$, puesto que de $f(x) = y$ y del hecho de que f posee una función inversa $f^{-1} : Y \to X$ (pues f es biunívoca de X sobre Y) concluimos que $f^{-1}(y) = x$, y como $g(y) = x$, se tiene que $g(y) = f^{-1}(y)$ para todo $y \in Y$, lo que concluye la prueba. QED

Notemos que en la segunda parte del enunciado de la proposición 3, los papeles de las funciones f y g son totalmente simétricos y, por consiguiente, es lícito concluir que $g : Y \to X$ es biunívoca de Y sobre X y tiene a $f : X \to Y$ como a su inversa.

Esta proposición muestra que las condiciones $f^{-1}f = I$ y $ff^{-1} = I$ caracterizan a la inversa f^{-1} de f en el sentido de que las condiciones $gf = I$ y $fg = I$ implican que f^{-1} existe y es igual a g. Adviértase que una sola de estas condiciones no basta para caracterizar a la inversa. Así, en el ejemplo 3 de este párrafo, se tiene $di = I$, pero $id \neq I$; en este ejemplo ni d ni i poseen inversa, pues aunque d sea sobre D, d no es inyectiva en D (dado que dos funciones distintas pueden tener la misma derivada) y, aunque i sea inyectiva en Y, i no es sobre D (dado que el conjunto de los valores de i está formado por los elementos de D que se anulan en el punto 0).

En lo que sigue tendremos varias oportunidades de emplear la segunda parte de la proposición 3, a fin de establecer una "identidad" entre conceptos distintos, tal como la que existe entre los conceptos de relación de equivalencia y de partición, etc.

Proposición 4. Sean $f : X \to Y$ y $g : Y \to Z$ dos funciones. Si f y g son inyectivas en X e Y, respectivamente, entonces gf es inyectiva en X. Análogamente, si f y g son sobre Y y Z respectivamente, entonces gf es sobre Z.

Demostración. Supongamos que f y g son inyectivas. Si $x, x' \in X$, $x \neq x'$, entonces $f(x) \neq f(x')$, dado que f es inyectiva en X. Se sigue que $g\{f(x)\} \neq g\{f(x')\}$, pues g es inyectiva en Y, o sea $(gf)(x) \neq (fg)(x')$, lo que establece la inyectividad de gf en X.

Supongamos ahora que f y g son sobre o suryectivas. Dado $z \in Z$ existe por lo menos un $y \in Y$ tal que $g(y) = z$, dado que g es sobre Z. Obtenido este $y \in Y$ existe por lo menos un $x \in X$ tal que $f(x) = y$, dado que f es sobre Y. Luego $g\{f(x)\} = z$, o sea, $(gf)(x) = z$, lo que implica que gf es sobre Z. QED

Proposición 5. Sean $f : X \to Y$ y $g : Y \to Z$ dos funciones. Si $f^{-1} : Y \to X$ y $g^{-1} : Z \to Y$ existen, entonces existe $(gf)^{-1} : Z \to X$ y $(gf)^{-1} = f^{-1}g^{-1}$.

Demostración. La inversa de una función existe sólo en el caso en que ésta sea biunívoca sobre (§ 5). En virtud de la proposición anterior vemos que la existencia de f^{-1} y g^{-1} implica la existencia de $(gf)^{-1}$. Además de esto,

$$x = (gf)^{-1}(z) \Leftrightarrow (gf)(x) = z \Leftrightarrow g\{f(x)\} = z \Leftrightarrow f(x) = g^{-1}(z) \Leftrightarrow$$
$$\Leftrightarrow x = f^{-1}\{g^{-1}(z)\} \Leftrightarrow x = (f^{-1}g^{-1})(z)$$

lo que prueba que $(gf)^{-1} = f^{-1}g^{-1}$. QED

Ejercicios

1) Sea $f : X \to Y$ una función. Para todo $A \subset X$, se tiene

$$A \subset f^{-1}\{f(A)\}.$$

Para que la siguiente igualdad valga para todo A

$$A = f^{-1}\{f(A)\}$$

es necesario y suficiente que f sea inyectiva en X. Análogamente,

$$f\{f^{-1}(B)\} \subset B$$

para cualquier $B \subset Y$. Para que sea válida la siguiente igualdad cualquiera que sea B,

$$f\{f^{-1}(B)\} = B$$

es necesario y suficiente que f sea sobre Y.

2) Sean X, Y y Z tres conjuntos.

a) Dadas dos funciones $g : Y \to Z$ y $h : X \to Z$, para que exista por lo menos una función $f : X \to Y$, tal que $h = gf$, es necesario y suficiente que

$$h(X) \subset g(Y).$$

Para que f sea única es necesario y suficiente que f sea inyectiva en Y.

b) Dadas dos funciones $f : X \to Y$ y $h : Y \to Z$, para que exista por lo menos una función $g : Y \to Z$, tal que $h = gf$, es necesario y suficiente que

$$f(x) = f(x') \Rightarrow h(x) = h(x'),$$

esto es, si los valores de f en dos puntos de X son iguales, entonces los valores de h en los mismos puntos también son iguales. Para que g sea única es necesario y suficiente que f sea sobre Y.

3) Sea $f : X \to X$ una función. Definamos $f^0 = I$, $f^1 = f$, $f^2 = ff$, y de un modo más general, $f^n = ff^{n-1} (n = 1, 2, \ldots)$, donde $I : X \to X$ es la función identidad de X. Si existe un entero $n \geq 2$ tal que $f^n = I$, entonces f es una permutación de X y $f^{-1} = f^{n-1}$. En particular, para que $f : X \to X$ sea una involución, es necesario y suficiente que $f^2 = I$.

4) Sean $f : X \to Y$ y $g : Y \to Z$ dos funciones y $gf : X \to Z$ su producto. Cualesquiera que sean $A \subset X$ y $C \subset Z$, se tiene

$$(gf)(A) = g\{f(A)\}, \quad (gf)^{-1}(C) = f^{-1}\{g^{-1}(C)\}.$$

§ 7. RELACIONES DE EQUIVALENCIA

En matemática elemental se encuentran de modo natural varios ejemplos importantes de *"relaciones binarias"*, esto es, de relaciones entre dos elementos cualesquiera de un mismo conjunto tomados en un cierto orden, que se distinguen por gozar de las propiedades reflexiva, simétrica y transitiva. Mencionemos algunos casos que son familiares al lector.

Ejemplo 1. Consideremos un conjunto E. La relación de igualdad $x = y$, donde x, $y \in E$, es una relación binaria en E, esto es, una relación entre elementos x e y de E considerados en este orden, que goza de las siguientes propiedades:

$$x = x,$$

$$x = y \Rightarrow y = x,$$

$$x = y,\ y = z \Rightarrow x = z.$$

Ejemplo 2. Consideremos un plano euclideano P y el conjunto R de las rectas de P. La relación de paralelismo $x \mathbin{//} y$, donde x, $y \in R$, es una relación binaria en R tal que

$$x // x,$$

$$x // y \Rightarrow y // x,$$

$$x // y,\ y // z \Rightarrow x // z.$$

Ejemplo 3. Consideremos un plano euclideano P y el conjunto S de los segmentos orientados de P. La relación de equipolencia $x \simeq y$, donde x, $y \in S$ (que se define como válida en los dos casos siguientes: 1) x e y no son nulos y tienen la misma dirección, el mismo sentido y la misma longitud; 2) x e y son nulos) constituye una relación binaria en S que posee las siguientes propiedades:

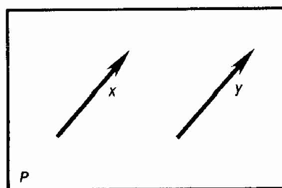


Fig. 15

$$x \simeq x,$$

$$x \simeq y \Rightarrow y \simeq x,$$

$$x \simeq y,\ y \simeq z \Rightarrow x \simeq z.$$

Además de esta relación en S, se acostumbra considerar también la relación de equipolencia restringida (definida como antes, salvo que, en el caso 1, se sustituye el requisito de que x e y tengan la misma dirección por la condición de que x e y tengan el mismo punto de apoyo), la cual goza también de las tres propiedades anteriores. Estas dos relaciones son indispensables, como se sabe, en la presentación de los conceptos de vector libre y vector deslizante.

Ejemplo 4. Consideremos la relación de congruencia en el conjunto **Z** de los números enteros racionales, ya estudiada en aritmética. Dos enteros x, $y \in \mathbf{Z}$ son congruentes módulo p, donde p es un entero natural, cuando $x - y$ es un múltiplo de p. Escribimos, entonces, $x \equiv y$ (mód. p). Si consideramos un p fijo, la congruencia módulo p constituye una relación binaria en **Z**, la cual, como es sabido, goza de las propiedades siguientes:

$$x \equiv x \text{ (mód. } p\text{)},$$

$$x \equiv y \text{ (mód. } p\text{)} \Rightarrow y \equiv x \text{ (mód. } p\text{)},$$

$$x \equiv y \text{ (mód. } p\text{)}, \; y \equiv z \text{ (mód. } p\text{)} \Rightarrow x \equiv z \text{ (mód. } p\text{)}.$$

De esta manera queda definida una relación binaria en **Z** para cada p. La congruencia módulo 0 es la propia relación de igualdad. Dos elementos cualesquiera de **Z** son siempre congruentes módulo 1.

Los ejemplos que se acaban de mostrar justificarían la introducción de la noción general de relación de equivalencia, aun si no existieran también otros temas en matemática donde las consideraciones que haremos en torno de esta noción encuentran aplicación. Consideremos un conjunto E. Una *relación de equivalencia* en E es una relación binaria en E que goza de las propiedades reflexiva, simétrica y transitiva. En general, se hará uso del símbolo "$\sim$" y se escribirá $x \sim y$ (leer "x equivale a y") para indicar que los elementos x, $y \in E$ guardan entre sí la relación considerada. Las propiedades de la relación de equivalencia son las siguientes:

e^1. $x \sim x$,

e^2. $x \sim y \Rightarrow y \sim x$,

e^3. $x \sim y$, $y \sim z \Rightarrow x \sim z$.

En circunstancias especiales, como en los ejemplos indicados arriba, se usan otros símbolos, tales como $\equiv$, $\simeq$, R, etc., y otras denominaciones, tales como congruencia, equipolencia, etc., para denotar una relación de equivalencia.

La noción de relación de equivalencia nunca debe ser disociada de la noción de partición, pues, como se verá más adelante, existe una conexión simple, pero importante, entre las dos. Una *partición* de un conjunto E es una colección P de conjuntos, cada uno de los cuales recibe el nombre de *componente* de la partición, tales que:

p^1. toda componente es un subconjunto no vacío de E;

p^2. todo elemento de E pertenece a una, y sólo una, componente.

Nótese que p^2 implica que dos componentes o son disjuntas o coinciden.

Con el fin de aclarar de inmediato esta noción se mencionan los ejemplos siguientes.

Ejemplo 5. Consideremos los conjuntos E, A_1, A_2, . . . , A_n ($n \geq 1$) que satisfacen las siguientes propiedades:

a. $A_i \neq \phi$ ($1 \leq i \leq n$); b. $E = A_1 \cup A_2 \cup \ldots A_n$; c. $A_i \cap A_j = \phi$ ($1 \leq i, j \leq n$; $i \neq j$).

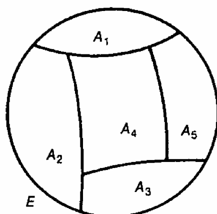


Fig. 16

Estos conjuntos A_1, A_2, . . . , A_n constituyen una partición de E en n componentes, pues por a. y b. todo A_i es un subconjunto no vacío de E, y por b. y c. todo punto de E pertenece a un, y sólo a un, A_i. Así, si consideráramos al conjunto de los enteros racionales y estuviéramos interesados en los varios residuos que la división por un $p \geq 1$ fijo pudiera dejar, nos inclinaríamos a agrupar en un mismo conjunto $\mathbf{Z}_i$ a los enteros de la forma $kp + i$, o sea, a los que dejan residuo i en la división por p ($0 \leq i \leq p - 1$). Los conjuntos $\mathbf{Z}_0$, $\mathbf{Z}_1$, . . . , $\mathbf{Z}_{p-1}$ constituyen una partición de $\mathbf{Z}$ en p elementos.

Ejemplo 6. Consideremos un plano euclideano P y una dirección fija d en P. La colección R_d de todas las rectas de P que tiene dirección d constituye una partición de P en una infinidad de componentes. En efecto, cada una de tales rectas es una parte no vacía de P. Además de esto, todo punto de P pertenece a una, y sólo a una, recta de dirección d.

Pasemos ahora a establecer la conexión que existe entre los conceptos de relación de equivalencia y de partición.

Proposición 1. Dada una partición P de un conjunto E, si $x \sim y$ se define por la condición de que los elementos x e y de E pertenezcan a la misma componente de P, se obtiene una relación de equivalencia en E.

Demostración. Por definición, escribiremos $x \sim y$ para indicar que existe una componente A de la partición tal que $x \in A$, $y \in A$. Ahora, dado $x \in E$, por p^2 existe una componente A de la partición tal que $x \in A$ y, entonces, $x \in A$, $x \in A$ muestra que $x \sim x$, lo que prueba e^1. Supongamos, ahora, que $x \sim y$, esto es, que exista una componente A tal que $x \in A$, $y \in A$. Como, entonces, $y \in A$, $x \in A$, vemos que $y \sim x$ y, por consiguiente, se satisface e^2. Finalmente, supongamos que $x \sim y$ e $y \sim z$. Como $x \sim y$ existe una componente A tal que $x \in A$, $y \in A$, y como $y \sim z$ existe otra componente A' tal que $y \in A'$, $z \in A'$. Si se observa que $y \in A$, $y \in A'$, a partir de p^2 concluimos que $A = A'$. Luego $z \in A' = A$ y, entonces, $x \in A$, $z \in A$ implica $x \sim z$, lo que prueba e^3. QED

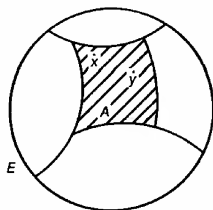


Fig. 17

De este modo queda establecido que toda partición de un conjunto determina una relación de equivalencia en el mismo, la cual se dice que

es *asociada* a la partición. Así, en el caso de la partición de **Z** indicada en el ejemplo 5, la relación de equivalencia asociada es precisamente la congruencia módulo p. En efecto, $x \sim y$ significa, por definición, que x e y pertenecen a un mismo $\mathbf{Z}_i$, o sea que x e y tienen el mismo residuo i cuando son divididos por p, o lo que es lo mismo, que $x - y$ es un múltiplo de p, esto es $x \equiv y$ (mód. p).

Antes de enunciar la proposición 2, que es una especie de recíproca de la que acabamos de establecer, vamos a introducir el concepto de clase de equivalencia. Consideremos un conjunto E y una relación de equivalencia en E. Llamamos *clase de equivalencia* de un elemento $x \in E$ al conjunto de los elementos $y \in E$ tales que $x \sim y$. Usaremos siempre las notaciones $[x]$ o $\bar{x}$ para representar la clase de equivalencia de x. Notemos que $x \in [x]$ pues, por e^1, $x \sim x$. Notemos también que $x \sim y$ si, y sólo si, $[x] = [y]$. En efecto, supongamos que $x \sim y$. Si $z \in [x]$, entonces $x \sim z$. Ya que $y \sim x$ y $x \sim z$, se tiene que $y \sim z$, de donde se sigue que $z \in [y]$, lo que prueba $[x] \subset [y]$. De forma análoga, $[y] \subset [x]$, de donde $[x] = [y]$. Recíprocamente, si $[x] = [y]$, como $y \in [y]$, vemos que $y \in [x]$, de donde se tiene $x \sim y$.

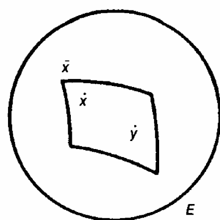


Fig. 18

Proposición 2. Dados un conjunto E y una relación de equivalencia en E, sus clases de equivalencia constituyen las componentes de una partición de E.

Demostración. Toda clase de equivalencia $[x]$ es, por definición, un subconjunto de E, el cual, además de esto, no es vacío; esto último resulta de que $x \in [x]$. Luego, p^1 queda satisfecha, Esta misma relación $x \in [x]$ muestra que todo elemento x de E pertenece a, por lo menos, una clase de equivalencia $[x]$. Sólo queda demostrar que cualquier clase de equivalencia $[y]$ que contenga a x debe ser igual a $[x]$. Ahora, $x \in [y]$ significa que $y \sim x$, lo que equivale a $[x] = [y]$. Luego, p^2 queda satisfecha. QED

De este modo queda establecido que toda relación de equivalencia en un conjunto determina una partición del mismo, que se dice *asociada* a la relación. Así, si pensamos en la congruencia módulo $p \geq 1$ como una relación de equivalencia en **Z**, la partición asociada a esta relación es precisamente la descrita en el ejemplo 5. En efecto, $x \equiv y$ (mód. p) es sinónimo de que x e y tienen el mismo residuo cuando se los divide por p; o sea la clase de equivalencia de x es el conjunto de los y que tienen el mismo residuo que x en esta división. Como los residuos posibles son $0, 1, \ldots, p - 1$, vemos que las clases de equivalencia posibles son $\mathbf{Z}_0, \mathbf{Z}_1, \ldots, \mathbf{Z}_{p-1}$.

El caso del ejemplo 2 da lugar a observaciones más interesantes. En cursos elementales de geometría se define la dirección de una recta x en un plano euclideano P como *lo que tienen de común* con x todas

las rectas del plano que le son paralelas. Es claro, sin embargo, que la expresión "lo que tienen de común" no pasa de ser un recurso de lenguaje que puede ser usado con éxito, en este caso y en algunos otros análogos, para transmitir una idea intuitiva como la de dirección. Empero, si tuviéramos más apego a la preocupación de formular los varios conceptos matemáticos en términos de las nociones de conjunto y función, se vuelve necesario definir dirección de otro modo. ¿Qué hay de común a dos rectas paralelas? La relación de paralelismo es, por sobre todo, una relación de equivalencia. Se dijo ya que en una relación de equivalencia dos elementos son equivalentes si, y sólo si, ellos determinan la misma clase de equivalencia. En particular, dos rectas de P son paralelas si, y sólo si, mediante la relación de paralelismo, determinan la misma clase en el conjunto R de las rectas de P. Luego, lo que dos rectas paralelas tienen en común es su clase de equivalencia. Por tanto, dentro del espíritu de economía de los conceptos primitivos de la matemática, se define la dirección de una recta x como su clase de equivalencia $[x]$ en R: la dirección de una recta es, entonces, un cierto conjunto.

Las observaciones hechas arriba pueden aplicarse también al ejemplo 3 de este párrafo. En vez entonces de definir un vector libre como *lo que hay de común* entre un segmento orientado y todos los que le son equipolentes, es más adecuado definirlo como la clase de equivalencia en el conjunto S de todos los segmentos orientados, clase ésta relativa a la relación de equipolencia. El vector libre de un segmento orientado x es, entonces, la clase $[x]$ que x determina. Comentarios análogos se aplican a la relación de equipolencia restringida a los vectores deslizantes.

Ejercicios

1) Dado el conjunto E, sean $P(E)$ el conjunto de las partes de E y $R(E)$ el conjunto de las relaciones de equivalencia en E. Las proposiciones 1 y 2 del § 7 definen dos aplicaciones $r : P(E) \to R(E)$ y $e : R(E) \to P(E)$. Probar que $pr = I$ y $rp = I$, esto es r y p son una inversa de la otra.

2) Sea $p(n)$ el número de las particiones posibles de un conjunto finito con n elementos (o también, el número de las relaciones de equivalencia posibles en tal conjunto). Hallar un método de cálculo por recurrencia para $p(n)$.

§ 8. ESPACIOS COCIENTES

Uno de los modos más usuales de definir una relación de equivalencia en un conjunto E, cuando E es el dominio de una cierta función, consiste en considerar dos puntos de E como equivalentes siempre y cuando la función asuma el mismo valor en ambos puntos. En efecto:

Proposición 1. Dada una función $f : E \to F$, si definimos, dados x, $y \in E$, $x \sim y$ cuando $f(x) = f(y)$, obtenemos una relación de equivalencia en E.

Demostración. Se tiene $x \sim x$, pues $f(x) = f(x)$. Si $x \sim y$, esto es, si $f(x) = f(y)$, entonces $f(y) = f(x)$, y en tal caso $y \sim x$. Por último, si

$x \sim y$ e $y \sim z$, esto es, si $f(x) = f(y)$ y $f(y) = f(z)$, entonces, $f(x) = f(z)$, y en tal caso $x \sim z$. Se sigue pues que la relación definida en E es de equivalencia. QED

Toda función determina, entonces, en su dominio, una relación de equivalencia, que se dice asociada a la función. Por ejemplo, si consideramos el conjunto $\mathbf{Z}$, un entero fijo $p \geq 1$ y una función $r_p : \mathbf{Z} \to \mathbf{N}$ que a cada $x \in \mathbf{Z}$ asocia su residuo $r_p(x)$ cuando es dividido por p, entonces la relación de equivalencia definida por $x \sim y$, si $r_p(x) = r_p(y)$, es precisamente la congruencia módulo p. Otro ejemplo: si consideramos la función de $\mathbf{R}$ en $\mathbf{R}$ definida por $x \mapsto x^2$, la relación de equivalencia $x \sim y$ en $\mathbf{R}$, dada por $x^2 = y^2$, es aquella en la cual cada $x \neq 0$ es equivalente sólo a x y a $-x$, y $x = 0$ es equivalente sólo a 0. Tercer ejemplo: si consideramos la operación de derivación (pág. 12), la relación de equivalencia $f \sim g$ en D definida por $f' = g'$ es aquélla por la cual dos funciones son equivalentes cuando difieren por una constante.

Cabe ahora naturalmente preguntarse sobre la validez de una proposición recíproca a la anterior. Más explícitamente, dados un conjunto E y una relación de equivalencia en E, ¿se puede siempre encontrar una cierta función $f : E \to F$, cuyo dominio sea E y cuya relación de equivalencia asociada sea exactamente la dada? La noción bastante sencilla de espacio cociente, que se introducirá con el objeto de mostrar que tal recíproca es verdadera, encontrará su aplicación en varios puntos importantes de la matemática, en especial en las construcciones destinadas a establecer la existencia de determinados tipos de sistemas.

Examinemos un conjunto E, en el cual está dada una relación de equivalencia R. Cada punto x de E determina una clase de equivalencia. Si se hace variar x en E y se consideran todas las clases de equivalencia que así se obtienen, se tendrá, por la proposición 2 del párrafo precedente, una cierta partición de E. Se denominará *espacio cociente* de E por la relación de equivalencia R al conjunto cuyos elementos son las clases de equivalencia de E. Este espacio cociente se representará por E/R. Los elementos de E/R son, entonces, ciertas partes de E, a saber: las partes de E que aparecen como clases de equivalencia de los elementos de E; en otros términos, $E/R \subset \mathcal{P}(E)$. Por ejemplo, en el caso de la figura 19, si consideramos la relación de equivalencia R que da lugar a la partición de E en las cuatro componentes indicadas, el espacio cociente E/R consistirá de cuatro elementos, a saber: los conjuntos A_1, A_2, A_3 y A_4. De modo análogo, el espacio cociente de $\mathbf{Z}$ por la relación de congruencia módulo $p \geq 1$ fijo consiste de p elementos, a saber: los conjuntos $\mathbf{Z}_0, \mathbf{Z}_1, \ldots, \mathbf{Z}_{p-1}$, pues éstos constituyen las varias componentes de la partición asociada a la congruencia módulo p (pág. 27). El espacio cociente del conjunto S de los segmentos orientados por la relación de equipolencia (pág. 24) es el conjunto de los vectores libres (pág. 28) del plano.

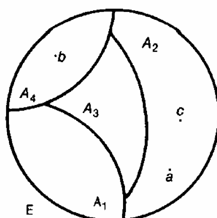


Fig. 19

Es importante cuando se considera el espacio cociente, que se piense en la llamada *proyección natural* de E en E/R. Todo elemento $x \in E$

pertenece a una, y sólo a una, clase de equivalencia $\bar{x}$. Ahora bien, por la definición de espacio cociente, $\bar{x}$ es un elemento del mismo. Si consideramos la correspondencia de E en E/R que a cada $x \in E$ asocia su clase de equivalencia $\bar{x} \in E/R$, tenemos definida una función $\pi : E \to E/R$, la cual denominamos proyección natural de E en el espacio cociente. Notemos, explícitamente, que

$$\pi(x) = \bar{x} \text{ para todo } x \in E.$$

Por ejemplo, en el caso de la figura 19, tenemos

$$\pi(a) = A_3, \quad \pi(b) = A_4, \quad \pi(c) = A_2, \text{ etc.};$$

análogamente, en el caso de la congruencia módulo $p = 3$ en $\mathbf{Z}$, tenemos

$$\pi(6) = \mathbf{Z}_0, \quad \pi(8) = \mathbf{Z}_2, \quad \pi(4) = \mathbf{Z}_1, \text{ etc.}$$

Tras presentar los conceptos de espacio cociente y proyección natural, cabe ahora presentar como recíproca a la proposición 1 la siguiente:

Proposición 2. Dados un conjunto E y una relación de equivalencia R en E, la proyección natural $\pi : E \to E/R$ es una función de E sobre E/R, que determina en E precisamente la relación de equivalencia R.

Demostración. Vamos a probar que π es sobre E/R. Todo elemento $A \in E/R$ es, por definición, la clase de equivalencia $\bar{x}$ de algún elemento $x \in E$, o séa $A = \bar{x} = \pi(x)$, como deseábamos. Además, por la definición de relación de equivalencia asociada a una función, dos elementos $x, y \in E$ son considerados como equivalentes por la relación de equivalencia en E asociada a π, si $\pi(x) = \pi(y)$, o sea, $\bar{x} = \bar{y}$, lo que significa (pág. 29) que x e y son equivalentes según R. Esto muestra que la relación de equivalencia determinada en E por la función π coincide con la relación R dada. QED

Vamos ahora a establecer una tercera proposición que, en cierto sentido, relaciona las situaciones descritas en las dos proposiciones anteriores y muestra cómo mediante una correspondencia biunívoca i, bien determinada entre F y E/R, se puede pasar de $f : E \to F$ a $\pi : E \to E/R$. Conviene recalcar que en la proposición siguiente la función f es sobre.

Proposición 3. Dada una función $f : E \to F$ de E sobre F, consideremos la relación de equivalencia R que f determina en E y la proyección natural $\pi : E \to E/R$. Existe, entonces, una, y sólo una, función $i : F \to E/R$ tal que $\pi = if$. Tal función i es biunívoca de F sobre E/R y, para todo $t \in F$, se tiene $i(t) = f^{-1}(t)$.

Demostración. Comencemos probando que si $x \in E$, $t \in F$ y $t = f(x)$, entonces $\bar{x} = f^{-1}(t)$. En efecto,

$$y \in \bar{x} \Leftrightarrow x \sim y \Leftrightarrow f(x) = f(y) \Leftrightarrow t = f(y) \Leftrightarrow y \in f^{-1}(t)$$

lo que establece $\bar{x} = f^{-1}(t)$. Luego, para todo $t \in F$, se tiene $f^{-1}(t) \in \in E/R$; en efecto, f es sobre F, entonces existe por lo menos un $x \in E$ tal que $t = f(x)$, de donde $f^{-1}(t) = \bar{x} \in E/R$. Así, es evidente que podemos definir una función $i : F \to E/R$ escribiendo $i(t) = f^{-1}(t)$, para $t \in \in F$. Esta función es sobre E/R.

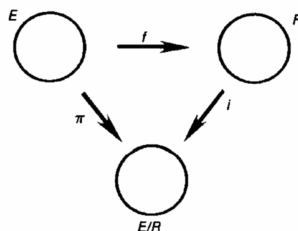


Fig. 20

De hecho, todo elemento $A \in E/R$ es una clase de equivalencia $\bar{x}$ para algún $x \in E$. Por consiguiente, si ponemos $t = f(x)$, vemos que $A = \bar{x} = = f^{-1}(t) = i(t)$, como queríamos. Además, i es inyectiva en F, pues si $t,\ t' \in F,\ t \neq t'$, y tuviésemos $i(t) = i(t')$, esto es, $f^{-1}(t) = f^{-1}(t')$, usando el hecho de que f es sobre F y escogiendo un x en E tal que $f(x) = = t$, esto es

$$x \in f^{-1}(t) = f^{-1}(t'),$$

tendríamos también $f(x) = t'$, donde $t = t'$, lo que contradice la suposición.

Probemos ahora que $\pi = it$. Dado $x \in E$ cualquiera y escribiendo $t = f(x)$, se tiene

$$i\{f(x)\} = i(t) = f^{-1}(t) = \bar{x} = \pi(x)$$

como se deseaba.

Resta probar la unicidad de $i : F \to E/R$ tal que $\pi = if$. Para ello, consideremos una función $j : F \to E/R$ tal que también $\pi = jf$. Dado $t \in \in F$ cualquiera, existe por lo menos un $x \in E$, para el cual $t = f(x)$ y, entonces, recurriendo a $if = jf$, vemos que

$$i\{f(x)\} = j\{f(x)\} \text{ o } i(t) = j(t)$$

lo que prueba realmente la igualdad entre i y j, como se afirmó. QED

El proceso que consiste en pasar de un conjunto a su espacio cociente por medio de una relación de equivalencia recibe el nombre de proceso de *identificación*, ya que en virtud del mismo dos elementos pertenecientes a una misma clase de equivalencia pasan a ser identificados con un único punto del espacio cociente. Este proceso es muy útil en geometría. A modo de ejemplo, se recordará que es posible *visualizar* ciertas propiedades de un plano proyectivo cuando se lo supone como la superficie de la esfera en el espacio euclideano tridimensional, en la cual todo punto es identificado con el punto diametralmente

opuesto. De este modo, la esfera resulta dividida en clases de equivalencia que constan de dos puntos. El plano proyectivo es, entonces, el espacio cociente de la esfera en virtud de tal relación.

Ejercicios

1) Sean $f : E \to F$ y $f' : E' \to F'$ dos funciones sobre F y F' que definen la misma relación de equivalencia en E. Existe, entonces, una, y sólo una, función $t : F \to F'$ tal que $f' = tf$. Tal función t es biunívoca de F sobre F'. Además de esto, si $t' : F' \to F$ fuese una función tal que $f = t'f'$, entonces t e t' son una la inversa de la otra.

2) Sean E y F dos conjuntos, en cada uno de los cuales está dada una relación de equivalencia. Si una función $f : E \to F$ es tal que

$$x \sim y \text{ en } E \text{ implica } f(x) \sim f(y) \text{ en } F,$$

entonces existe una, y sólo una, función $g : E/R \to F/S$, donde R y S son las relaciónes de equivalencia dadas en E y F, tal que

$$\pi_1 f = g\pi_2,$$

donde $\pi_1 : F \to F/S$ y $\pi_2 : E \to E/R$ son las proyecciones naturales ($\pi_1 = \pi_2$).

§ 9. PRODUCTOS CARTESIANOS FINITOS

A Descartes se debe la introducción y el empleo sistemático de los sistemas de coordenadas en el estudio de cuestiones geométricas y el consecuente florecimiento de la geometría analítica. Aunque recientemente ésta ha sido reducida a proporciones adecuadas a fin de ceder lugar a los métodos *invariantes*, la idea de Descartes dejó una huella definitiva en matemática, a saber: el concepto de *producto cartesiano*. Como veremos, este concepto en su forma general constituye una instancia más de la noción de función.

Si consideramos un plano euclideano P y un sistema de coordenadas Oxy en P, todo punto de P determina y queda completamente determinado por sus coordenadas x e y; de ahí la costumbre de sustituir un punto geométrico por el par ordenado (x, y) de sus coordenadas y decir, por ejemplo, el punto (x, y), etc. Se acostumbra, asimismo, llevar más alla tal identificación entre los objetos geométricos y sus coordenadas para, invirtiendo los respectivos papeles, emplear estas últimas en la definición de aquéllos. Así, en análisis matemático, después de definir el conjunto $\mathbf{R}$ de los números reales a partir del conjunto $\mathbf{Q}$ de los números racionales, por los métodos de Dedekind o de Cantor, se define un punto del plano euclideano $\mathbf{R}^2$ como un par ordenado (x, y) de números reales x e y, siendo entonces $\mathbf{R}^2$ el conjunto de tales pares ordenados. En forma más general,

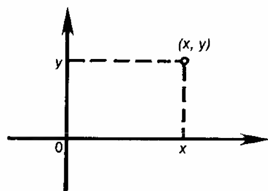


Fig. 21

se define también un punto del espacio euclideano $\mathbf{R}^n$ de dimensión n como una secuencia ordenada $(x_1, x_2, \ldots, x_n)$ de n números reales $x_1, x_2, \ldots, x_n$, y, entonces, $\mathbf{R}^n$ es el conjunto de tales secuencias. La noción de producto cartesiano de un número finito de factores se infiere de forma similar a lo que se acaba de indicar para $\mathbf{R}^2$ y $\mathbf{R}^n$ sin la restricción de que las coordenadas tengan necesariamente un significado numérico.

Empecemos considerando dos conjuntos E y F. Llamamos *producto cartesiano*, o simplemente *producto*, de E por F, representado por $E \times F$, al conjunto cuyos elementos son los pares ordenados (x, y), esto es, los pares formados cada uno de ellos por los elementos x e y considerados en este orden, donde $x \in E$ e $y \in F$. Los conjuntos E y F se denominan *factores*; x es la *primera coordenada* del punto (x, y) e y es su *segunda coordenada*. Por consiguiente, de acuerdo con esta terminología, el plano $\mathbf{R}^2$ es el producto cartesiano $\mathbf{R} \times \mathbf{R}$ de $\mathbf{R}$ por sí mismo. De manera general, se llama *cuadrado cartesiano*, o simplemente *cuadrado*, del conjunto E, y se representa por E^2, al producto $E \times E$. Nótese que, en virtud del sentido atribuido a la noción de igualdad (pág. 2) dos pares ordenados (x, y) y (x', y') son iguales si, y sólo si, $x = x'$ e $y = y'$.

Existen dos funciones

$$\pi_E : E \times F \to E \text{ y } \pi_F : E \times F \to F$$

asociadas al producto $E \times F$, que se definen del siguiente modo. La primera, que recibe el nombre de *proyección en* E, es la correspondencia que a todo punto (x, y) de $E \times F$ asocia su coordenada x de E, o sea,

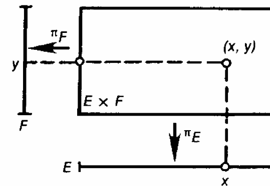


Fig. 22

$$\pi_E(x, y) = x;$$

la segunda es la correspondencia que a todo (x, y) asocia su coordenada y,

$$\pi_F(x, y) = y,$$

de ahí el nombre de proyección en F.

La generalización al caso de un número finito de factores es inmediata. El *producto cartesiano* de los conjuntos factores $E_1, E_2, \ldots, E_n$, en este orden, es el conjunto que se representa por

$$E_1 \times E_2 \times \ldots \times E_n \text{ o } \Pi_{i=1}^n E_i \text{ o } \Pi_i E_i$$

y cuyos elementos son las secuencias ordenadas $(x_1, x_2, \ldots, x_n)$, donde $x_1 \in E_1$, $x_2 \in E_2, \ldots, x_n \in E_n$. Cada x_i recibe el nombre de i-ésima coordenada del punto respectivo. Para ser concisos, se acostumbra representar el punto $(x_1, x_2, \ldots, x_n)$ por (x_i), donde se sobrentiende que i toma los valores $1, 2, \ldots, n$; o, aún más sencillamente, por la misma letra x que designa sus coordenadas:

$$x = (x_i) = (x_1, x_2, \dots, x_n).$$

En el caso de que todos los factores sean iguales a un mismo conjunto E, el producto $E \times E \times \dots \times E$ (n veces) se denomina *n-ésima potencia cartesiana* de E, y se representa, entonces, por E^n. Se aplica también a la noción de igualdad de secuencias la observación hecha en el caso $n = 2$.

Además, asociadas al producto $\Pi_i E_i$, existen n operaciones de proyección

$$\pi_i : E_1 \times E_2 \times \dots \times E_n \rightarrow E_i \ (i = 1, 2, \dots, n),$$

donde π_i es una versión simplificada de la notación π_{E_i}, cada una de las cuales se define como la correspondencia que a cada punto del producto asocia su i-ésima coordenada:

$$\pi_i(x) = x_i, \quad \text{si} \quad x = (x_1, x_2, \dots, x_n).$$

Ejemplo 1. En la práctica aparecen con frecuencia las siguientes potencias cartesianas:

$\mathbf{R}^n$ = espacio real euclideano de dimensión n;

$\mathbf{C}^n$ = espacio complejo euclideano de dimensión n;

$\mathbf{Z}^n$ = retículo de dimensión n (Fig. 23);

$\mathbf{T}^n$ = toro de dimensión n, donde $\mathbf{T}$ es el conjunto de los números complejos de módulo 1. La denominación "toro" tiene el siguiente origen. En el espacio euclideano tridimensional consideremos un eje E,

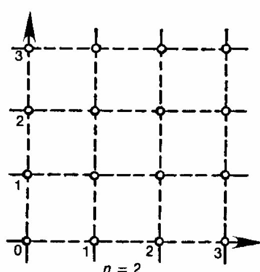


Fig. 23

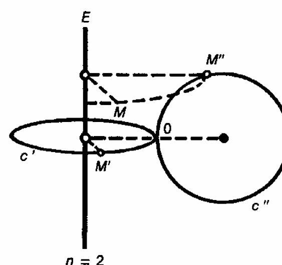


Fig. 24

una circunferencia C' con centro en E, cuyo plano sea perpendicular a E, y una circunferencia C'' en un plano que contiene a E y que intersecte a C' en el punto O (Fig. 24). Mediante una rotación completa de C'' alrededor de E se obtiene la superficie S del bien conocido toro. Ahora bien, todo punto (M', M'') del producto cartesiano $C' \times C''$, donde $M' \in C'$ y $M'' \in C''$, determina un punto M de S de la siguiente forma: M es el punto

al cual M'' es llevado por la rotación de C'' alrededor de E, cuando se hace coincidir O con M'. Es claro que de este modo se establece una correspondencia biunívoca entre el producto cartesiano $C' \times C''$ y el toro S. Conviene advertir además que si se representan a los números complejos por los puntos del plano de Gauss (esto es, $x + yi$, donde x, $y \in \mathbf{R}$), $\mathbf{T}$ se representará por la circunferencia con centro en el origen y de radio unidad. De ahí el nombre de toro dado a $\mathbf{T} \times \mathbf{T}$ y, por extensión, a $\mathbf{T}^n$. El propio $\mathbf{T}$ recibe el nombre de toro unidimensional.

La noción de producto cartesiano permite que se formule el concepto del *gráfico* de una función $f : E \to F$. Se da este nombre al conjunto G de los puntos del producto $E \times F$ que son de la forma $(x, f(x))$, donde x varía en E. El gráfico de una función de E en F es, entonces, un cierto subconjunto del producto $E \times F$. Por ejemplo, el gráfico de la transformación identidad $I : E \to E$ de un conjunto E es el subconjunto Δ del cuadrado E^2 formado por los puntos de la forma (x, x), donde x varía en E. Este conjunto se denomina *diagonal* del cuadrado E^2.

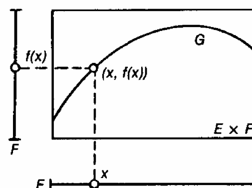


Fig. 25

Se llama *función de n variables* a toda función $f : E \to F$, cuyo dominio E es un producto cartesiano de n conjuntos, esto es,

$$E = E_1 \times E_2 \times \ldots \times E_n.$$

Si $x = (x_1, x_2, \ldots, x_n)$ designa un punto de E, el valor $f(x)$ de f en x también será indicado por $f(x_1, x_2, \ldots, x_n)$, como sugiere la notación fx (pág. 11) y no por $f((x_1, x_2, \ldots, x_n))$ como sugeriría la notación $f(x)$. Conviene de todas maneras no perder de vista que una función de n variables no es más que una función de un elemento x de un producto cartesiano de n factores. Es éste el sentido en el cual toda función de varias variables $x_1, x_2, \ldots, x_n$ debe ser pensada como una función de una variable $x = (x_1, x_2, \ldots, x_n)$. Un caso de funciones de dos variables que va a surgir con frecuencia es el de las *composiciones binarias* en un conjunto E. Se da tal nombre a toda función definida en E^2 con valores en E. Así, por ejemplo, la correspondencia $(x, y) \mapsto x + y$ que a todo par ordenado de números reales asocia su suma, define una ley de composición binaria en el conjunto $\mathbf{R}$ de los números reales. Análogamente, $(x, y) \mapsto xy$ define otra composición binaria en $\mathbf{R}$. Otro ejemplo: consideremos un conjunto E y sea $\mathcal{P}(E)$ el conjunto de las partes de E. La correspondencia $(X, Y) \mapsto X \cup Y$, que a todo par ordenado de subconjuntos de E asocia su unión constituye una composición binaria en $\mathcal{P}(E)$. Lo mismo se aplica a la correspondencia $(X, Y) \mapsto X \cap Y$. Tercer ejemplo: consideremos un conjunto E y sea E^E el conjunto de las funciones de E en E. Si $f : E \to E$ y $g : E \to E$ designan dos de tales funciones, se puede siempre formar la función compuesta $gf : E \to E$ y, entonces, la correspondencia $(f, g) \mapsto gf$ constituye una composición binaria en E^E.

Las observaciones precedentes se refieren al caso en que el dominio de una función es un producto cartesiano. Tratemos ahora el caso de funciones definidas en un conjunto E con valores en otro conjunto

$$F = F_1 \times F_2 \times \ldots \times F_n$$

que se presenta como un producto de n factores $F_1, F_2, \ldots, F_n$. Si

$$f_1 : E \to F_1,\ f_2 : E \to F_2, \ldots,\ f_n : E \to F_n$$

designan n funciones, se puede, a partir de éstas, construir una función $f : E \to F$, a saber, la función que a cada punto x de E asocia el punto

$$(f_1(x), f_2(x), \ldots, f_n(x)) \in F;$$

o sea, se define f por

$$f(x) = (f_1(x), f_2(x), \ldots, f_n(x)).$$

Tal función f se denomina producto cartesiano de las n funciones dadas y, entonces, se escribe

$$f = f_1 \times f_2 \times \ldots \times f_n$$

Cada f_i recibe el nombre de *componente* de f. Nótese que las componentes de f se expresan por medio de esta función del siguiente modo:

$$f_i = \pi_i f \quad (i = 1, 2, \ldots, n),$$

donde $\pi_i : F_1 \times F_2 \times \ldots \times F_n \to F_i$ indica la proyección en F_i. En efecto,

$$\pi_i\{f(x)\} = \pi_i(f_1(x), f_2(x), \ldots, f_n(x)) = f_i(x)$$

para todo $x \in E$, como se quería. Recíprocamente, toda función $f : E \to F$ es el producto cartesiano de n componentes. En efecto, definamos $f_i = \pi_i f (i = 1, 2, \ldots, n)$. Notemos que si $y = (y_1, y_2, \ldots, y_n)$ designa un punto cualquiera de F, entonces $\pi_i(y) = y_i$ y, por tanto,

$$y = (\pi_1(y), \pi_2(y), \ldots, \pi_n(y)),$$

donde

$$f(x) = (\pi_1 f(x), \pi_2 f(x), \ldots, \pi_n f(x)) = (f_1(x), f_2(x), \ldots, f_n(x))$$

para todo $x \in E$, lo que prueba $f = f_1 \times f_2 \times \ldots \times f_n$. En particular, sean $f : \mathbf{R} \to \mathbf{R}$ y $g : \mathbf{R} \to \mathbf{R}$ dos funciones reales continuas de variable real. Su producto cartesiano es la función $f \times g$ de $\mathbf{R}$ en $\mathbf{R}^2$ definida por

$$t \mapsto (f(t), g(t)),\ t \in \mathbf{R}.$$

Tal función $f \times g$ constituye lo que se acostumbra denominar la representación paramétrica de una curva continua en el plano $\mathbf{R}^2$. Así, si consideramos las funciones $t \mapsto \cos t$ y $t \mapsto \operatorname{sen} t$, su producto cartesiano $t \mapsto (\cos t, \operatorname{sen} t)$ representa la parametrización de una circunferencia en $\mathbf{R}^2$.

Observación. Aunque se ha mencionado ya (págs. 1 y 28) que los conjuntos y las funciones debieran ser los dos elementos básicos en

términos de los cuales se procuraría formular las demás nociones, hemos fallado ya a este respecto en dos puntos, a saber: en la presentación del concepto de relación binaria (pág. 24) y en el de par ordenado (pág. 32) y, más generalmente, en el de secuencia ordenada (pág. 33). Es que, por ser las expresiones de relación binaria, par ordenado y secuencia ordenada muy expresivas, nos pareció aconsejable dejar hasta ahora los comentarios que siguen.

Dar una secuencia ordenada $x = (x_1, x_2, \ldots, x_n)$ consiste en indicar su primer término x_1, su segundo término $x_2, \ldots$, su n-ésimo término x_n. Si se representa por I_n el conjunto finito constituido por los enteros 1, 2, ..., n, la secuencia x puede ser interpretada como una función definida en el conjunto I_n, función que puede llamarse x y cuyo valor en el punto $i \in I_n$, que se indica por x_i y no por $x(i)$, se denomina i-ésima coordenada de x. En particular, un par ordenado puede definirse como una función cuyo dominio es el conjunto constituido por los enteros 1 y 2, y cuyos valores de la función en los puntos 1 y 2 son el primero y el segundo elemento, respectivamente, del par ordenado.

1. → $\cdot x_1$

2. → $\cdot x_2$

. .

. .

. .

n. → $\cdot x_n$

Fig. 26

Habiendo formulado los conceptos de par ordenado y secuencia ordenada en términos de funciones, podemos definir un producto cartesiano finito.

$$\Pi_i E_i = E_1 \times E_2 \times \ldots \times E_n$$

como el conjunto de todas las funciones x definidas en I_n, cada una de las cuales está sujeta a la condición de que su valor x_i en el punto $i \in I_n$ pertenezca al conjunto E_i. Un producto cartesiano finito entonces pasa a ser un conjunto de funciones. Es importante que se señale desde ya que tal presentación de un producto cartesiano finito como un cier-

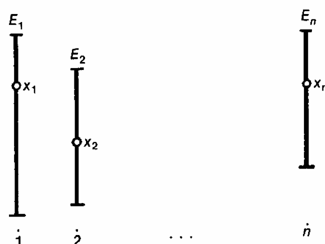


Fig. 27

to conjunto de funciones no es fruto de una obsesión por los conjuntos y funciones, sino más bien, al contrario, constituye una necesidad, pues, como se verá más adelante, es en esta forma que el concepto de producto cartesiano arbitrario --finito o infinito-- puede formularse de forma natural.

Pasemos ahora al caso de las relaciones binarias. Consideremos el ejemplo 2 de la página 24. La relación de paralelismo en el conjunto R de las rectas del plano P determina un subconjunto G del cuadrado R^2, denominado *gráfico* de la relación, a saber: el conjunto de los pares ordenados (x, y), donde $x, y \in R$, y tales que $x // y$. Tal conjunto G, a su vez, determina la relación de paralelismo, pues dos rectas $x, y \in R$ guardan entre sí la relación de paralelismo si, y sólo si, $(x, y) \in G$. En vez, entonces, de pensar en que G queda definido por la relación de paralelismo, podemos invertir los papeles, y suponiendo que G es dado, definir la relación de paralelismo como el propio conjunto G. Lo que acabamos de indicar en el caso de la relación de paralelismo se aplica también a cualquier otra relación binaria. Por consiguiente, se puede definir una relación binaria en un conjunto E como un subconjunto G del cuadrado E^2 y decir que dos elementos $x, y \in E$ guardan entre sí la relación considerada cuando $(x, y) \in G$. Una relación de equivalencia en el conjunto E pasa entonces a ser definida como un subconjunto G del cuadrado E^2 tal que

e^1. $(x, x) \in G$,

e^2. $(x, y) \in G \Rightarrow (y, x) \in G$,

e^3. $(x, y) \in G,\ (y, z) \in G \Rightarrow (x, z) \in G$,

donde $x, y, z \in E$ (véase la pág. 25). Más aún, además de las relaciones binarias en un conjunto E, esto es, de las relaciones entre dos elementos arbitrarios x, y en E, en las aplicaciones aparecen relaciones entre elementos de dos conjuntos E y F, esto es, relaciones entre elementos arbitrarios $x \in E$, $y \in F$ (el caso de las relaciones binarias en E corresponde a $E = F$). Ejemplo: la relación de paralelismo entre el conjunto de las rectas del espacio euclideano tridimensional y el conjunto de los planos de este espacio. Por las razones que se acaban de mencionar, se puede definir una *relación* entre dos conjuntos E y F como un subconjunto de $E \times F$.

Las consideraciones precedentes muestran cómo se pueden introducir los pares ordenados a partir de las funciones. Inversamente, el concepto de función puede formularse por medio de conjuntos y pares ordenados. En efecto, admitiendo estos dos últimos conceptos como primitivos, el producto cartesiano $E \times F$ de dos conjuntos E y F tiene sentido (pág. 33). Cabe entonces definir una función de E en F como el subconjunto G del producto $E \times F$ que goza de la siguiente propiedad: para todo $x \in E$ existe uno, y sólo un, $y \in F$ tal que $(x, y) \in G$. La idea es clara: estamos definiendo una función por medio de su gráfico en vez de definir éste a partir de aquélla, como se hizo en la página 35. Resumiendo: se ha visto que es indiferente tomar como conceptos primitivos los de conjunto y función, o los de conjunto y par ordenado. Los autores que adoptan la segunda alternativa explican su preferencia afirmando que los conceptos primitivos de matemática son los de conjunto y orden.

Ejercicios

1) El producto cartesiano $E \times F$ de dos conjuntos finitos de m y n elementos es también finito y tiene mn elementos. Similarmente, para $E_1 \times E_2 \times \ldots \times E_n$.

2) En álgebra elemental se establecen propiedades de suma, producto y potencia:

$$xy = yx, \quad z^{x+y} = z^x z^y, \quad (yz)^x = y^x z^x, \quad (z^y)^x = z^{yx}.$$

Si X e Y designan dos conjuntos, escribamos $X \cong Y$ para indicar que existe por los menos una correspondencia biunívoca entre X e Y. Demostrar que

$$X \times Y \simeq Y \times X, \quad Z^{X+Y} \simeq Z^X \times Z^Y, \quad (Y \times Z)^X \simeq Y^X \times Z^X, \quad (Z^Y)^X \simeq Z^{X \times Y},$$

donde X, Y y Z son conjuntos y, en la segunda relación, X e Y se asumen disjuntos y $X + Y$ representa $X \cup Y$.

3) Sean E, F y G tres conjuntos. Si A es una parte de $E \times F$ y B es una parte de $F \times G$, definamos el producto AB como la parte de $E \times G$ formada por los pares (x, z) para cada uno de los cuales existe por lo menos un y tal que $(x, y) \in A$ y $(y, x) \in B$. Demostrar que este producto es asociativo y que las diagonales (pág. 35) actúan como unidad con respecto a tal producto.

Además de esto, si A es una parte de $E \times F$, definamos A^{-1} como la parte de $F \times E$ formada por los pares (y, x) tales que $(x, y) \in A$. Demostrar que $(BA)^{-1} = A^{-1}B^{-1}$, donde el producto se entiende en el sentido mencionado arriba.

4) Sea E un conjunto. Demostrar que una relación de equivalencia en E puede ser definida como una parte G del cuadrado E^2 tal que

$$\Delta \subset G, \quad G^{-1} = G, \quad GG \subset G,$$

donde Δ es la diagonal de E^2 y G^{-1}, GG deben entenderse en el sentido del ejercicio anterior.

§ 10. INDICES

El presente párrafo sólo tiene por objeto mencionar un detalle de notación y terminología.

Considérese una función $f : X \to Y$. En muchas situaciones es costumbre representar el valor de la función f en el punto $x \in X$ por f_x y no por $f(x)$, como se ha hecho hasta ahora. Un caso típico es aquel en que el dominio X de la función considerada está constituida por números enteros. Cuando X es finito y consiste de los enteros 1, 2, ..., n, la función $f : X \to Y$ recibe el nombre de *secuencia* (véase la observación al final del párrafo precedente), y entonces se representa por

$$(f_1, f_2, \ldots, f_n) \quad \text{o} \quad (f_x)_{x=1}^{n}$$

Cuando X es infinito y está constituido por los enteros 1, 2, ..., n, ..., la función $f : X \to Y$ se denomina *sucesión* y se representa por

$$(f_1, f_2, \ldots, f_n, \ldots) \quad \text{o} \quad (f_x)_{x=1}^{\infty}.$$

Además de los dos casos que se acaban de mencionar, existen otros en que se prefiere la notación f_x a $f(x)$ a fin de resaltar el papel meramente auxiliar o enumerativo de x. En tales situaciones se acostumbra dar el nombre de *índice* al elemento x que varía en X, y al dominio X de la función se le llama el *conjunto de índices*. La propia función $f : X \to Y$ recibe, entonces, el nombre de *familia* de elementos de F, y pasa a ser representada por

$$(f_x)_{x \in X}$$

o por otras notaciones como (f_x) cuando no cabe duda de cuáles son los conjuntos X e Y. Por ejemplo, si a cada número real x le asociamos el intervalo

$$A_x = [-x^2, x^2 + 1]$$

de la recta **R**, tendremos así una familia $(A_x)_{x=-\infty}^{x=+\infty}$ de intervalos de **R**. En general, se acostumbra emplear las letras ι, j, α, λ, etc. para las variables escritas en la posición de índice.

§ 11. UNIONES E INTERSECCIONES ARBITRARIAS

Las nociones de unión e intersección presentadas en el § 3, páginas 5-10, para el caso de un número finito de conjuntos, pueden extenderse sin dificultad al caso de un número infinito de conjuntos.

Consideremos un conjunto E y sea $\mathcal{A}$ una colección de subconjuntos de E. Se llama unión de los elementos de $\mathcal{A}$, o más brevemente unión de $\mathcal{A}$, al conjunto de todos los elementos $x \in E$ que pertenecen a por lo menos uno de los conjuntos que constituyen la colección $\mathcal{A}$. De forma análoga, la intersección de los miembros de $\mathcal{A}$, o más brevemente la intersección de $\mathcal{A}$, es la colección de los elementos $x \in E$ que pertenecen a todos los conjuntos que constituyen $\mathcal{A}$. La unión e intersección de $\mathcal{A}$ se representan por

$$\cup \mathcal{A}, \quad \cap \mathcal{A}.$$

Por ejemplo, dados un punto V y una curva C en el espacio euclideano tridimensional, la superficie cónica de vértice V y directriz C es la unión de las semirrectas con origen en V y que se apoyan en C. De igual modo, dados tres puntos en un plano euclideano, la intersección de los círculos del plano que contienen a estos puntos en su interior es el triángulo determinado por los mismos.

Pasemos ahora a considerar un conjunto E y una familia $(X_\iota)_{\iota \in I}$ de subconjuntos de E, o sea una correspondencia que a todo elemento ι de un cierto conjunto I asocia una parte X_ι de E. Se llama unión de la familia $(X_\iota)_{\iota \in I}$ a la colección de los elementos $x \in E$ que pertenecen a uno de los conjuntos X_ι por lo menos. La intersección se define como la colección de todos los elementos $x \in E$ que pertenecen a todos los conjuntos X_ι. Tales unión e intersección se representan por

$$\cup_{\iota \in I} X_\iota \quad \text{y} \quad \cap_{\iota \in I} X_\iota$$

o por notaciones similares, tales como $\bigcup_\iota X_\iota$ y $\bigcap_\iota X_\iota$, siempre y cuando no hubiese duda sobre cuál es el conjunto I de los índices. Por ejemplo, en el caso de la familia (A_x), mencionada al final de la sección anterior, la unión $\bigcup_x A_x$ es el propio conjunto R y la intersección $\bigcap_x A_x$ es el intervalo $[0, 1]$.

Cabe naturalmente la siguiente pregunta ¿cuál es la diferencia, en las definiciones de arriba, entre los casos de una colección de conjuntos y una familia de conjuntos? Tal diferencia, aunque pequeña, consiste en el hecho de que cuando consideramos una familia $(X_\iota)_{\iota \in I}$ de conjuntos, no excluimos el hecho de que a dos índices distintos ι_1 e ι_2 pueda corresponder el mismo conjunto, esto es, $X_{\iota_1} = X_{\iota_2}$ o, como se acostumbra a decir, una familia no excluye la repetición de un mismo conjunto entre sus miembros; al contrario, en el caso de una colección todos los elementos son distintos. Esta diferencia, sin embargo, no tiene mayor importancia en la formación de uniones e intersección en virtud de que $X \cup X = X$ y $X \cap X = X$.

Ejercicios

1) Establecer las siguientes leyes distributivas:

$$X \cap \bigcup_{\jmath \in J} Y_\jmath = \bigcup_{\jmath \in J} (X \cap Y_\jmath)$$

$$\bigcap_{\iota \in I} X_\iota \cap \bigcup_{\jmath \in J} Y_\jmath = \bigcup_{(\iota,\jmath) \in I \times J} (X_\iota \cap Y_\jmath)$$

y sus duales que se obtienen permutando los signos de unión e intersección.

2) Establecer la siguiente ley de dualidad:

$$\complement\{\bigcup_{\iota \in I} X_\iota\} = \bigcap_{\iota \in I} \complement X_\iota,$$

y su dual que se obtiene permutando los signos de unión e intersección, donde los complementos se toman en relación con un mismo conjunto E.

3) Establecer las siguientes propiedades:

$\bigcup_{\iota \in I} X_\iota \subset Y$ equivale a $X_\iota \subset Y$ para todo $\iota \in I$,

$X \subset \bigcap_{\jmath \in J} Y_\jmath$ equivale a $X \subset Y_\jmath$ para todo $\jmath \in J$.

4) Si $J \subset I$, mostrar que

$$\bigcup_{\iota \in J} X_\iota \subset \bigcup_{\iota \in I} X_\iota, \quad \bigcap_{\iota \in I} X_\iota \subset \bigcap_{\iota \in J} X_\iota.$$

5) Si $X_\iota \subset Y_\iota$, para todo $\iota \in I$, establecer

$$\bigcup_{\iota \in I} X_\iota \subset \bigcup_{\iota \in I} Y_\iota,$$

y su dual que se obtiene sustituyendo $\cup$ por $\cap$.

6) Demostrar la ley asociativa

$$\bigcup_{\iota \in I} X_\iota = \bigcup_{t \in T}\{\bigcup_{\iota \in J_t} X_\iota\},$$

y su dual, donde $(J_t)_{t \in T}$ designa una familia de subconjuntos de I tal que

$$\bigcup_{t \in T} (J_t) = I.$$

7) Sean $X_1, X_2, \ldots, X_n, \ldots$ una sucesión de conjuntos. Se llama *límite superior* de la sucesión a la colección de los puntos que pertenecen a un número infinito de miembros de la sucesión $\{X_n\}$; llamamos *límite inferior* a la colección de puntos que dejan de pertenecer sólo a un número finito de miembros de la sucesión. Establecer las siguientes fórmulas

$$\lim_{n \to \infty} \sup X_n = \bigcap_{m=1}^{\infty} \{\bigcup_{n=m}^{\infty} X_n\},$$

$$\lim_{n \to \infty} \inf X_n = \bigcup_{m=1}^{\infty} \{\bigcap_{n=m}^{\infty} X_n\}.$$

8) Sea $f : E \to F$ una función. Para cualquier familia $(X_\iota)_{\iota \in I}$ de subconjuntos de E se tiene

$$f(\bigcup_\iota X_\iota) = \bigcup_\iota f(X_\iota), \quad f(\bigcap_\iota X_\iota) \subset \bigcap_\iota f(X_\iota).$$

En la segunda relación se cumple la igualdad si, y sólo si, la función es inyectiva. Análogamente, para toda familia $(Y_\jmath)_{\jmath \in J}$ de subconjuntos de F, se tiene

$$f^{-1}(\bigcup_\iota X_\iota) = \bigcup_\iota f^{-1}(X_\iota), \quad f^{-1}(\bigcap_\iota X_\iota) = \bigcap_\iota f^{-1}(X_\iota).$$

§ 12. PRODUCTOS CARTESIANOS ARBITRARIOS

El concepto de producto cartesiano de un número finito de factores puede ser ampliado al caso de una familia arbitraria de factores del siguiente modo.

Consideremos una familia $(E_\iota)_{\iota \in I}$ de conjuntos, o sea una correspondencia que a cada elemento $\iota \in I$ asocia un conjunto E_ι. Inspirándonos en la definición de producto cartesiano finito indicada en la página 37, fijemos nuestra atención en las funciones x definidas en el conjunto I, cada una de las cuales está sujeta a la condición de que su valor x_ι en el punto $\iota \in I$ debe pertenecer a E_ι, para todo $\iota \in I$. La colección de tales funciones se denominará *producto cartesiano de la familia* de conjuntos $(E_\iota)_{\iota \in I}$ y se representará por

$$\Pi_{\iota \in I} E_\iota$$

o por notaciones similares, tales como $\Pi_\iota E_\iota$. Cada E_ι recibe el nombre de ι-ésimo factor del producto. El valor x_ι de la función x en $\iota \in I$ es conocido como *ι-ésima coordenada* del punto x del producto. Se acostumbra representar un elemento x del producto por $(x_\iota)_{\iota \in I}$, o simplemente por (x_ι), siempre que convenga mencionar sus coordenadas x_ι. La igualdad de dos elementos del producto debe entenderse en el sentido ya mencionado para las funciones (pág. 11), o sea dos elementos (x_ι) e (y_ι) del producto son iguales si $x_\iota = y_\iota$, para todo $\iota \in I$.

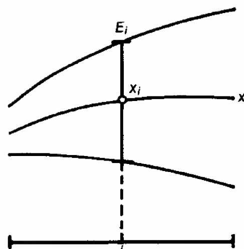


Fig. 28

En el caso particular en que todos los factores E_ι sean iguales a un mismo conjunto E, el producto cartesiano $\Pi_{\iota\in I}E_\iota$ se denomina *potencia cartesiana* I de E. Recordando la definición de producto, vemos que un elemento x del producto es una función definida en I, cuyo valor x_ι en $\iota \in I$ debe pertenecer a $E_\iota = E$; o sea un elemento del producto es simplemente una función definida en I con valores en E. Ahora bien, la notación E^I fue ya atribuida a la colección de las funciones de I en E (pág. 12). Resumiendo,

$$\Pi_{\iota\in I}E_\iota = E^I \text{ si } E_\iota = E \text{ para todo } \iota \in I.$$

Por ejemplo, el conjunto de las funciones reales de variable real definidas en un intervalo $[a, b]$ es un producto cartesiano de un número infinito de factores, cada uno de los cuales corresponde a un número del intervalo y es igual al propio conjunto **R** de los números reales.

La *proyección* del producto $\Pi_{\jmath\in I}E_\jmath$ en un factor E_ι es la correspondencia

$$\pi_\iota : \Pi_{\jmath\in I}E_\jmath \to E_\iota$$

que a cada punto x del producto asocia su ι-ésima coordenada x_ι, o sea,

$$\pi_\iota(x) = x_\iota, \ \iota \in I.$$

A toda función f definida en un producto cartesiano $E = \Pi_{\iota\in I}E_\iota$, con valores en otro conjunto F, se le da el nombre de *función de varias variables*. Conforme el conjunto I de los índices sea finito de n elementos o infinito, f recibe el nombre de *función de* n o *de infinitas variables*.

Dado un conjunto E y un producto cartesiano $F = \Pi_{\iota\in I}F_\iota$, toda familia de funciones $(f_\iota)_{\iota\in I}$

$$f_\iota : E \to F_\iota, \ \iota \in I,$$

permite que se defina una nueva función $f : E \to F$, a saber: la función que a cada punto $x \in E$ asocia el punto

$$f(x) = (f_\iota(x))_{\iota\in I} \in \Pi_{\iota\in I}F_\iota = F.$$

Tal función se denomina *producto cartesiano* de las funciones dadas y se representa por

$$f = \Pi_{\iota\in I} f_\iota.$$

Las componentes f_ι se expresan por medio de f, del siguiente modo:

$$f_\iota = \pi_\iota f,$$

donde $\pi_i : \Pi_{j\in I} F_j \to F_i$ es la proyección en F_i y, recíprocamente, uno puede verificar que toda función $f : E \to F$ es el producto cartesiano de las funciones $(f_i)_{i\in I}$ definidas por $f_i = \pi_i f$.

A fin de evitar malentendidos, conviene mencionar que, en el caso del producto $E_1 \times E_2 \times \ldots \times E_n$, la relación de orden usual $1 < 2 < \ldots < n$ entre los enteros permite escribir ordenadamente

$$(x_1, x_2, \ldots, x_n)$$

las varias coordenadas de un punto del producto. La misma observación cabe en el caso de un producto enumerable

$$\Pi_{i=1}^{\infty} E_i = E_1 \times E_2 \times \ldots \times E_n X \ldots,$$

cuyos elementos son sucesiones

$$(x_1, x_2, \ldots, x_n, \ldots)$$

de puntos $x_1 \in E_1$, $x_2 \in E_2, \ldots, x_n \in E_n, \ldots$. Sin embargo, en el caso general de $\Pi_{i\in I} E_i$, no se puede suponer ninguna relación de orden en el conjunto I de los índices y la actitud correcta consiste en pensar en los elementos del producto exclusivamente como funciones, sin insistir en ningún orden de las varias coordenadas de los elementos de este producto.

2

GRUPOS

Los grupos tuvieron su origen en la teoría de sustituciones debida en parte a los trabajos de Lagrange. Sin embargo, el verdadero iniciador de este capítulo del álgebra fue Galois. El desarrollo de la teoría de grupos estaba en ese entonces condicionado a sus aplicaciones a la teoría de las ecuaciones algebraicas. Más tarde, los trabajos de Sophus Lie mostraron la importancia de los grupos en ciertos aspectos de las ecuaciones diferenciales y abrieron camino a la teoría de los llamados grupos de Lie, y las ideas de Felix Klein, relacionadas con la conveniencia de considerar a la geometría como el estudio de las propiedades invariantes por determinados grupos de transformaciones, ampliaron el campo de aplicación del concepto de grupo. En su forma axiomática, la noción de grupo fue introducida en el siglo pasado por Cayley y abarca dos aspectos: los grupos aditivos y los multiplicativos. Los primeros constituyen (excepto por cambios de notación) un caso particular de los segundos. Por motivos didácticos, al comienzo de este capítulo mencionaremos explícitamente los diversos aspectos de la teoría en sus versiones aditiva y multiplicativa. Luego, a partir de un cierto punto, nos limitaremos a formular los conceptos y resultados en una de las dos notaciones, dejando la otra a cargo del lector.

§ 1. GRUPOS ADITIVOS

En matemática elemental se encuentran varios casos de conjuntos cuyos elementos pueden combinarse algebraicamente por medio de una operación de adición, de modo que algunas de las propiedades usuales sean satisfechas, a saber: la conmutatividad, la asociatividad, la existencia del cero y la existencia del simétrico (inverso aditivo). Mencionaremos algunos ejemplos:

Ejemplo 1. La operación de adición usual en el conjunto **Z** de los números enteros racionales es una función de $\mathbf{Z} \times \mathbf{Z}$ en **Z** que, a cada par ordenado (x, y), donde $x, y \in \mathbf{Z}$, asocia un elemento $x + y \in \mathbf{Z}$, denominado suma de x e y. Tal adición es conmutativa y asociativa:

$$x + y = y + x, \quad x + (y + z) = (x + y) + z.$$

Además, en **Z** existe un elemento cero, representado por 0, tal que:

$$x + 0 = x$$

para cualquier $x \in \mathbf{Z}$. Finalmente, a cada $x \in \mathbf{Z}$ corresponde un simétrico $-x \in \mathbf{Z}$ tal que:

$$x + (-x) = 0.$$

Análogamente, cada uno de los conjuntos **Q**, **R** y **C** (pág. 1) posee, con relación a la operación de adición usual, las propiedades que acaban de mencionarse respecto a Z.

Ejemplo 2. Consideremos un plano euclideano P, un punto O de P y el conjunto S_0 de los segmentos orientados de P con origen en O. Si x, $y \in S_0$, definamos $x + y$ de acuerdo con la regla usual del paralelogramo. La operación de adición así definida en S_0 es conmutativa y asociativa como se demuestra en el cálculo vectorial. El segmento orientado, cuyo origen y extremo son iguales al punto O, se representa por O y goza de la propiedad de que $x + O = x$ para todo $x \in S_0$. Finalmente, a cada $x \in S_0$ corresponde un simétrico $-x$, obtenido por simetría de x con respecto al punto O, tal que $x + (-x) = 0$. Así vemos que a pesar de que sus elementos están desprovistos del carácter numérico, el conjunto S_0 tiene, con respecto a la adición mencionada, un comportamiento idéntico al de los conjuntos del ejemplo precedente.

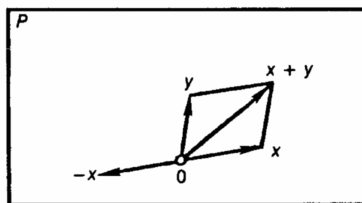


Fig. 29

Ejemplo 3. Consideremos el conjunto F de las funciones reales de variable real definidas en un intervalo $[a, b]$. Si f, $g \in F$, introduzcamos la suma $f + g$ como la función definida por

$$(f + g)(x) = f(x) + g(x)$$

para $x \in [a, b]$. O sea, $f + g$ es la función que en el punto x toma el valor $f(x) + g(x)$. La operación de adición así definida en F es conmutativa, pues

$$(f + g)(x) = f(x) + g(x),$$

$$(g + f)(x) = g(x) + f(x),$$

de donde $(f + g)(x) = (g + f)(x)$ para todo $x \in [a, b]$, o sea $f + g = g + f$. Análogamente,

$$[f + (g + h)](x) = f(x) + (g + h)(x) =$$

$$= f(x) + [g(x) + h(x)],$$

$$[(f + g) + h](x) = (f + g)(x) + h(x) =$$

$$= [f(x) + g(x)] + h(x),$$

de donde $[f + (g + h)]\ (x) = [(f + g) + h]\ (x)$ para cualquier $x \in [a, b]$, o sea $f + (g + h) = (f + g) + h$. Si representamos por 0 a la función idénticamente nula, esto es, la función cuyo valor en todo punto de $[a, b]$ es cero, es claro que $f + 0 = f$ para toda $f \in F$. Finalmente, a toda $f \in F$ le corresponde una función $-f$ definida por

$$(-f)\ (x) = -f(x),$$

o sea, $-f$ es la función que en el punto x toma el valor $-f(x)$, tal que $f + (-f) = 0$. La analogía entre este ejemplo y los anteriores es evidente. Conviene destacar que el hecho que las funciones en consideración hayan sido definidas en un intervalo $[a, b]$ no influye en estas consideraciones, que bien podrían repetirse para la colección de funciones definidas en un conjunto E y con valores en **R**.

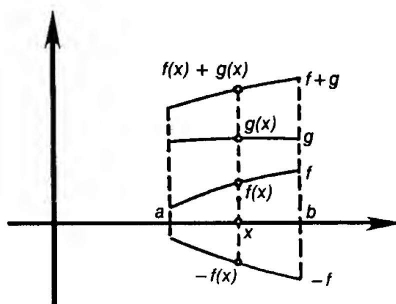


Fig. 30

La simple revisión de los ejemplos dados justifica plenamente la introducción del concepto de grupo aditivo.

Un *grupo aditivo* es un conjunto G, donde está dada una operación de adición que satisface las condiciones siguientes:

1. La operación de adición es una función definida en $G \times G$, con valores en G, que a cada par ordenado (x, y), donde $x, y \in G$, asocia un elemento $x + y \in G$, denominado *suma de los elementos* x e y.

2. La adición es *conmutativa*, esto es,

$$x + y = y + x, \qquad (x, y \in G).$$

3. La adición es *asociativa*

$$x + (y + z) = (x + y) + z, \qquad (x, y, z \in G).$$

4. En G existe un elemento, el *cero* de G (representado por 0), tal que

$$x + 0 = x \text{ para todo } x \in G.$$

5. A cada elemento $x \in G$ le corresponde un elemento denominado *simétrico* de x y representado por $-x$, tal que

$$x + (-x) = 0.$$

Los varios ejemplos presentados constituyen, pues, grupos aditivos con relación a las operaciones de adición mencionadas.

En virtud de la ley asociativa, podemos definir sin ambigüedad la suma de tres o más elementos. Así, por definición,

$$x + y + z = x + (y + z) = (x + y) + z.$$

Por el signo Σ se indicará en forma abreviada una suma

$$x_1 + x_2 + \ldots + x_n = \Sigma_{i=1}^{n} x_i.$$

Además, a semejanza de lo que sucede en álgebra elemental, las leyes conmutativa y asociativa permiten cierta libertad en la inversión del orden de los sumandos y en el empleo de paréntesis. El hecho de que se sumen elementos de un conjunto G en vez de números no introduce, desde este punto de vista, diferencia alguna. Lo que es esencial es que se tengan presentes los axiomas de los grupos aditivos arriba mencionados.

En un grupo aditivo G, el elemento cero está perfectamente determinado por la condición de que $x + 0 = x$ cualquiera que sea $x \in G$. En efecto, consideremos dos elementos 0 y 0' tales que

$$x + 0 = x, \qquad x + 0' = x$$

para todo $x \in G$. Si se escribe $x = 0'$ en la primera y $x = 0$ en la segunda, se obtiene $0' + 0 = 0'$ y $0 + 0' = 0$, de donde $0 = 0'$ por la ley conmutativa.

Además, el simétrico $-x$ de todo $x \in G$ está determinado por la condición de que su suma con x es cero. En efecto, sean $-x$ y $(-x)'$ dos elementos de G tales que

$$x + (-x) = 0, \qquad x + (-x)' = 0$$

Tenemos, entonces,

$$(-x)' = (-x)' + 0 = (-x)' + [x + (-x)] = [(-x)' + x] + (-x) =$$

$$= 0 + (-x) = -x,$$

como queríamos. La unicidad del simétrico puede interpretarse así:

$$x + y = 0 \Rightarrow y = -x.$$

La diferencia de dos elementos $x, y \in G$ se define por

$$x - y = x + (-y).$$

Esta diferencia es la única solución z de la igualdad $z + y = x$. En efecto, partiendo de $z + y = x$ y sumándole $-y$ obtenemos $(z + y) + (-y) = x + (-y)$, o $z + [y + (-y)] = x - y$ y, finalmente, $z + 0 = x - y$, lo que prueba que, si es que existe, la solución z está dada por la diferencia $x - y$. Es fácil pues verificar que $x - y$ es realmente la solución, ya que $(x - y) + y = [x + (-y)] + y = x + [(-y) + y] = x + 0 = x$.

La noción de *suma algebraica*, tal como $x - y + z$, tiene el mismo sentido que en álgebra elemental, y en este sentido se adoptan las convenciones usuales. Así, por ejemplo, $-x -y$ significa $(-x) + (-y)$ y no $-(x - y)$. Adviértase que $0 + 0 = 0$, por la propia definición de 0; de allí que $-0 = 0$ y que $0 - 0 = 0$.

Proposición 1. En todo grupo aditivo G, se tiene:

1. $-(-x) = x$,

2. $-(x + y) = -x -y$,

3. $x + y = x \Rightarrow y = 0$,

4. $x + y = x + z \Rightarrow y = z$.

Demostración. Dado $x \in G$, se tiene $x + (-x) = 0$ y $(-x) + x = 0$, lo que prueba que $-(-x) = x$ en virtud de la unicidad del simétrico.

Cabe notar ahora que $(x + y) + (-x-y) = (x + y) + [(-x) + (-y)] = [x + (-x)] + [y + (-y)] = 0 + 0 = 0$, lo que prueba que $-(x + y) = (-x-y)$. De un modo perfectamente análogo veríamos que

$$-(x - y) = y - x$$

y similarmente para otras sumas algebraicas.

Supongamos ahora que $x + y = x$. Entonces, $(-x) + (x + y) = (-x) + x$, luego $[(-x) + x] + y = 0$, o también $0 + y = 0$ y $y = 0$. Conviene hacer aquí una observación. El elemento 0 fue caracterizado por la condición $x + 0 = x$ para cualquier $x \in G$. Lo que se acaba de establecer muestra que el 0 es caracterizado también por el hecho de que la condición $x + 0 = x$ sea satisfecha por, al menos, un $x \in G$.

Consideremos, finalmente, la relación $x + y = x + z$. Tenemos que $(-x) + (x + y) = (-x) + (x + z)$, de donde $[(-x)+x] + y = [(-x)+x] + z$ o $0 + y = 0 + z$ y $y = z$. Esta última propiedad se denomina *ley de cancelación* y desempeña, como se verá, un papel importante en álgebra. QED

Todos los grupos aditivos mencionados en los ejemplos 1, 2 y 3 son *infinitos*, esto es, formados por una infinidad de elementos. Vamos ahora a dar un ejemplo fundamental de grupo aditivo *finito*, es decir, con un número finito de elementos.

Ejemplo 4. Dado un entero $p \geq 1$, consideremos el conjunto, representado por $\mathbf{Z}/p$, constituido por los enteros $0, 1, \ldots, p - 1$. Ad-

viértase desde ya que este conjunto no es un grupo aditivo con respecto a la operación de adición usual entre los enteros porque la suma de dos elementos de $\mathbf{Z}/p$ puede dejar de pertenecer a $\mathbf{Z}/p$. Entretanto, vamos a definir una operación de adición en $\mathbf{Z}/p$, con respecto a la cual este conjunto será un grupo aditivo. Si $x, y \in \mathbf{Z}/p$, definamos la *suma módulo* p de x e y como el residuo en la división por p de la suma usual $x + y$. Usaremos el símbolo $x \overset{p}{+} y$ para designar la suma módulo p de x e y. Como todo residuo de una división por p es igual a uno de los enteros $0, 1, \ldots, p - 1$, vemos que la adición módulo p en $\mathbf{Z}/p$ verifica el primero de los axiomas de los grupos aditivos. Es claro que la adición módulo p es conmutativa, pues $x \overset{p}{+} y$ e $y \overset{p}{+} x$ son, respectivamente, los residuos en la división por p de $x + y$ e $y + x$ y $x + y = y + x$. Verifiquemos ahora que la adición módulo p es asociativa, esto es

$$x \overset{p}{+} (y \overset{p}{+} z) = (x \overset{p}{+} y) \overset{p}{+} z$$

Para obtener $y \overset{p}{+} z$ hay que dividir $y + z$ entre p. Si se llama q al cociente y r al residuo de tal división, se tiene

$$y \overset{p}{+} z = r, \qquad y + z = pq + r.$$

Para calcular ahora $x \overset{p}{+} (y \overset{p}{+} z) = x \overset{p}{+} r$, debemos dividir $x + r$ entre p. Si se indica con q' el cociente y con r' el residuo, se tiene

$$x \overset{p}{+} (y \overset{p}{+} z) = r', \qquad x + r = pq' + r'.$$

De ahí resulta

$$(x + r) + (y + z) = (pq' + r') + (pq + r)$$

o

$$x + y + z = p(q + q') + r'.$$

El segundo miembro de la última igualdad es la suma de un múltiplo de p con un entero r', el cual, por ser el residuo de la división de $x + r$ entre p, sólo puede valer $0, 1, \ldots, p - 1$. Si recordamos la definición de residuo de una división, concluimos que r' es el residuo entre p de $x + y + z$, o sea

$$x \overset{p}{+} (y \overset{p}{+} z) = \text{residuo de } (x + y + z) \text{ entre } p$$

Un cálculo perfectamente análogo prueba que

$$(x \overset{p}{+} y) \overset{p}{+} z = \text{residuo de } (x + y + z) \text{ entre } p$$

lo que establece la ley asociativa de la adición módulo p. Notemos ahora que, para todo $x = 0, 1, \ldots, p - 1$, la suma $x \overset{p}{+} 0$ es el residuo de la división entre p de $x + 0 = x$. Como tal residuo es x, vemos que $x \overset{p}{+} 0 = x$, lo que prueba la existencia de un elemento cero en $\mathbf{Z}/p$, a saber, el propio 0. Finalmente, a todo $x \in \mathbf{Z}/p$ le corresponde un simétrico $\overset{p}{-}x \in \mathbf{Z}/p$ tal que $x \overset{p}{+} (\overset{p}{-}x) = 0$. En efecto, si $x > 0$ (esto es, $0 < x < p$ o $0 < p - x < p$), entonces $p - x$ es un elemento de $\mathbf{Z}/p$ y $x \overset{p}{+} (p - x)$ es el residuo de dividir $x + (p - x) = p$ entre p, o sea es igual a 0.

Si $x = 0$, $p - x = p$ no es un elemento de $\mathbf{Z}/p$. Notemos, entonces, que $0 \overset{p}{+} 0$ es el residuo de la división entre p de $0 + 0$, o sea que vale 0. Así, queda probado que el simétrico $\overset{p}{-}x$ de x módulo p existe en todos los casos y que

$$\overset{p}{-}x = \begin{cases} p - x & \text{si } x = 1, 2, \ldots, p - 1 \\ 0 & \text{si } x = 0 \end{cases}$$

$\mathbf{Z}/p$ constituye, pues, un grupo aditivo finito con relación a la adición módulo p. Se utilizó la notación $x \overset{p}{+} y$ para mayor claridad. Se acostumbra emplear la notación $x + y$ en cualquier grupo aditivo, en particular para $\mathbf{Z}/p$, y entonces es necesario tener presente si la suma está siendo calculada en el sentido usual o módulo p.

Ejercicios

1. Si el conjunto G posee uno, dos o tres elementos, entonces existe una, y sólo una, operación de adición con relación a la cual G es un grupo aditivo, salvo por permutaciones de sus elementos.

2. Sea G un conjunto y $(x, y) \mapsto x + y$ una operación conmutativa y asociativa de $G \times G$ en G tal que, cualesquiera que sean a, $b \in G$, la ecuación $a + x = b$ posee, por lo menos, una solución x. Mostrar que G es un grupo aditivo.

3. Sea G un conjunto y $(x, y) \mapsto x + y$ una operación conmutativa y asociativa de $G \times G$ en G tal que, cualesquiera que sean a, $b \in G$, la ecuación $a + x = b$ posee, como máximo, una solución x. Mostrar que G es un grupo aditivo.

§ 2. GRUPOS MULTIPLICATIVOS

Además de las operaciones de adición $(x, y) \mapsto x + y$ y de simetrización $x \mapsto -x$, también se consideran en álgebra elemental las operaciones de multiplicación $(x, y) \mapsto xy$ y de inversión $x \mapsto x^{-1}$ que gozan, entre otras, de las propiedades descritas en los ejemplos que siguen.

Ejemplo 1. La multiplicación usual puede ser efectuada entre números racionales cualesquiera. Como sólo tiene sentido tomar el inverso de números diferentes de cero y el producto de dos números racionales diferentes de cero también es diferente de cero, conviene restringir nuestra atención al conjunto $\mathbf{Q}^*$ de los números racionales no nulos. En $\mathbf{Q}^*$, la multiplicación usual es una función de $\mathbf{Q}^* \times \mathbf{Q}^*$ en $\mathbf{Q}^*$, que a cada par ordenado (x, y), donde x, $y \in \mathbf{Q}^*$, le asocia el producto $xy \in \mathbf{Q}^*$. Son válidas las leyes conmutativa y asociativa.

$$xy = yx, \qquad x(yz) = (xy)z.$$

Además, en $\mathbf{Q}^*$ existe un elemento unidad 1, tal que

$$x1 = x$$

para todo $x \in \mathbf{Q}^*$. Finalmente, a cada $x \in \mathbf{Q}^*$ le corresponde un inverso $x^{-1} \in \mathbf{Q}^*$ tal que

$$xx^{-1} = 1.$$

Lo mismo se repite para los conjuntos $\mathbf{R}^*$ y $\mathbf{C}^*$ de los números reales y de los números complejos no nulos con respecto a la multiplicación usual.

Ejemplo 2. En álgebra elemental, una permutación de $1, 2, \ldots, n$ se define como una secuencia constituida por dichos números escritos en algún orden determinado. También se introduce el concepto de producto de dos permutaciones. Para nuestros propósitos, es preferible (pág. 37) pensar en una permutación $(x_1, x_2, \ldots, x_n)$ de los enteros $1, 2, \ldots, n$ como una función definida en el conjunto I_n en el sentido indicado en la citada página y, entonces, las permutaciones de $1, 2, \ldots, n$ pasan a ser las aplicaciones biunívocas de I_n sobre sí mismo. Ese es el motivo por el cual definimos (pág. 16) una permutación de un conjunto E como una aplicación biunívoca de E sobre sí mismo. Generalizando las permutaciones del álgebra elemental, consideremos un conjunto E --finito o infinito-- y designemos con $E!$ el conjunto de las permutaciones de E, o sea el conjunto de las aplicaciones biunívocas de E sobre sí mismo. Si $f : E \to E$ y $g : E \to E$ fueran dos permutaciones de E, su producto $gf : E \to E$ es también una permutación de E, como resulta de la proposición 4, página 22. Así, hemos definido una multiplicación en el conjunto $E!$ que a cada par ordenado (g, f) de elementos de $E!$ asocia el producto $gf \in E!$ Tal multiplicación es asociativa

$$h(gf) = (hg)f,$$

como muestra la proposición 1, página 20. Ella, empero, no es necesariamente conmutativa. Por ejemplo, en el caso en que E esté constituido por los números 1, 2, y 3 y $f = \begin{pmatrix} 1, & 2, & 3 \\ 2, & 1, & 3 \end{pmatrix}$, $g = \begin{pmatrix} 1, & 2, & 3 \\ 1, & 3, & 2 \end{pmatrix}$, entonces $gf = \begin{pmatrix} 1, & 2, & 3 \\ 3, & 1, & 2 \end{pmatrix}$ y $fg = \begin{pmatrix} 1, & 2, & 3 \\ 2, & 3, & 1 \end{pmatrix}$; luego, $gf \neq fg$. Además, la transformación idéntica $I : E \to E$ es una permutación de E tal que $fI = If = f$, para toda $f \in E!$ (véase la proposición 2, pág. 20). Finalmente, a toda permutación $f : E \to E$ le corresponde una permutación inversa $f^{-1} : E \to E$ tal que $ff^{-1} = f^{-1}f = I$ (proposición 3, pág. 21). Así, vemos que el conjunto $E!$ de las permutaciones tiene, con relación a la multiplicación de permutaciones, un comportamiento análogo al de los conjuntos $\mathbf{Q}^*$, $\mathbf{R}^*$ y $\mathbf{C}^*$ con relación a la multiplicación de números, salvo en lo que respecta a la conmutatividad.

Antes de dar nuevos ejemplos de conjuntos en los que existe una multiplicación con las propiedades indicadas en los ejemplos citados, formularemos el concepto general de grupo multiplicativo.

Un *grupo multiplicativo* es un conjunto G en el cual se encuentra una multiplicación que satisface las propiedades siguientes:

1. La operación de multiplicación es una función definida en $G \times G$ con valores en G, que a cada par ordenado (x, y) en $G \times G$ le asocia un elemento $xy \in G$ denominado *producto* de los factores x e y.

2. La multiplicación es asociativa:

$$x(yz) = (xy)z \qquad (x, y, z \in G).$$

3. En G existe un elemento, denominado *unidad* de G y representado por e, tal que:

$$xe = ex = x \text{ para todo } x \in G.$$

4. A cada elemento $x \in G$ le corresponde un elemento denominado *inverso* de x y representado por x^{-1}, tal que:

$$xx^{-1} = x^{-1}x = e.$$

Conviene notar que la ley conmutativa no aparece en estos axiomas, pues nuestro deseo es definir la noción de grupo conmutativo de manera que englobe no sólo al ejemplo 1 sino también al 2. Cuando $xy = yx$, se dice que los elementos x e y *conmutan*. El grupo multiplicativo G se dice *conmutativo* o *abeliano* cuando $xy = yx$ para cualesquiera x e y en G. Por lo tanto, $\mathbf{Q}^*$, $\mathbf{R}^*$ y $\mathbf{C}^*$ constituyen grupos multiplicativos conmutativos con relación a la multiplicación usual. En $E!$ tenemos un ejemplo de grupo multiplicativo no necesariamente conmutativo: $E!$ se designa con el nombre de *grupo de las permutaciones* o *grupo simétrico* de E.

La ley asociativa nos permite definir, sin ambigüedad, el producto de tres o más factores. Así, por definición

$$xyz = x(yz) = (xy)z.$$

Como en álgebra elemental, haremos uso del signo Π para representar abreviadamente el producto

$$x_1x_2 \ldots x_n = \Pi_{i=1}^{n} x_i$$

y, siguiendo el ejemplo del caso aditivo, supondremos que el lector está familiarizado con el empleo (omisión e inserción) de paréntesis en los productos de tres o más factores. En el caso conmutativo, también podemos permutar dos factores cualesquiera de un producto. Nunca debemos olvidar, sin embargo, que el caso general no es tan sencillo.

En todo grupo multiplicativo, el elemento unidad está perfectamente determinado por las condiciones $xe = ex = x$, $x \in G$. En efecto, consideremos dos elementos $e, e' \in G$ tales que

$$xe = ex = x, \; xe' = e'x = x$$

Demostración. Como $x(x^{-1}) = (x^{-1})\ x = e$, vemos que x es el elemento de G que, multiplicado por x^{-1} de los dos modos posibles, produce la unidad. Luego, $x = (x^{-1})^{-1}$. Además, para llegar a esta conclusión, bastaría usar $xx^{-1} = e$ y la observación hecha al final de la demostración de la unicidad del inverso (pág. 54).

Notemos ahora que

$$(xy)\ (y^{-1}x^{-1}) = x(yy^{-1})x^{-1} = xex^{-1}\ xx^{-1} = e,$$

lo que basta para probar que $y^{-1}x^{-1} = (xy)^{-1}$, por la observación mencionada, sin que haya necesidad de verificar también que $(y^{-1}x^{-1})\ (xy) = e$.

Supongamos ahora $xy = x$. De ahí que $x^{-1}xy = x^{-1}x$, de donde $ey = e$ y, al final, $y = e$. Análogamente para la segunda parte del punto 3. Aquí cabe una observación semejante a la hecha en la página 49 para el caso similar de los grupos aditivos.

Finalmente, si $xy = xz$, entonces $x^{-1}xy = x^{-1}xz$, donde $ey = ez$, o sea $y = z$. La propiedad que acaba de ser establecida se designa con el nombre de *ley de cancelación a la izquierda*. La *ley de cancelación a la derecha*, expresada por la segunda parte del punto 4, se establece en forma análoga. QED

La noción de grupo multiplicativo puede ser formulada también de una manera más breve que la adoptada. En efecto:

Proposición 2. Un grupo multiplicativo puede ser definido también como un conjunto G que satisface las condiciones 1 y 2 de la página 53 y las condiciones 3' y 4' siguientes:

3'. En G existe un elemento *unidad a la derecha* e tal que

$$xe = x \text{ para todo } x \in G.$$

4'. A cada elemento $x \in G$ le corresponde un *inverso a la derecha* x^{-1} tal que

$$xx^{-1} = e.$$

Demostración. Está claro que si las condiciones 1, 2, 3 y 4 de la definición de grupo multiplicativo son satisfechas, entonces 1, 2, 3' y 4' también se verifican. Recíprocamente, partamos de estas últimas. Comencemos probando que el inverso a la derecha x^{-1} de todo elemento $x \in G$ es también un inverso a la izquierda, esto es que $x^{-1}x = e$. Por 4', todo elemento de G posee un inverso a la derecha. En particular, no sólo x posee un inverso a la derecha x^{-1} tal que $xx^{-1} = e$, sino que el elemento x^{-1} tiene un inverso a la derecha $(x^{-1})^{-1}$ tal que $x^{-1}(x^{-1})^{-1} = e$. Luego, por la ley asociativa, $xx^{-1} = e$ implica $x^{-1}xx^{-1} = x^{-1}e$, y por 3', obtenemos $x^{-1}xx^{-1} = x^{-1}$. De ahí que

$$x^{-1}xx^{-1}(x^{-1})^{-1} = x^{-1}(x^{-1})^{-1}$$

y aplicando $x^{-1}(x^{-1})^{-1} = e$ en los dos miembros, queda $x^{-1}xe = e$, de donde $x^{-1}x = e$ por 3', como queríamos. Mostremos ahora que la unidad a la derecha también actúa como unidad a la izquierda. En efecto, como $e = xx^{-1}$ se sigue que $ex = xx^{-1}x$, de donde $ex = xe$, pues $x^{-1}x = e$, como acabamos de ver y, para terminar, $ex = x$ por 3'. Vemos así que las condiciones 3 y 4 también son satisfechas. QED

La proposición anterior no dice que una unidad a la derecha sea necesariamente una unidad a la izquierda, o sea que la condición 3' individualmente implique 3; ni que todo inverso a la derecha sea un inverso a la izquierda, o sea que 4', por sí sola, implique 4. Lo que la proposición muestra es que, en presencia de la ley asociativa, el conjunto de las condiciones 3' y 4' implica 3 y 4. Es claro también que, además de la proposición anterior, vale una dual con respecto a las unidades y los inversos a la izquierda.

La única diferencia entre un grupo aditivo y un grupo multiplicativo conmutativo se reduce a una cuestión de notación. Más precisamente, supongamos que G sea un grupo aditivo, al cual está asociada una operación definida en $G \times G$ con valores en G. Si resolviésemos representar el resultado de esta operación sobre un par (x, y) no por $x + y$, como en la condición 1 de la página 47, sino por xy, entonces G pasaría a ser a las claras un grupo multiplicativo conmutativo: la conmutatividad y la asociatividad en la notación aditiva se convertirían en la conmutatividad y la asociatividad en la notación multiplicativa, el elemento de G que desempeña el papel de cero en la notación aditiva pasaría a desempeñar el papel de unidad multiplicativa y, por último, el simétrico de cada elemento de G se convertiría en su inverso. Al contrario, si G designase a un grupo multiplicativo conmutativo y resolviésemos representar por $x + y$ lo que estábamos designando por xy, es claro que G pasaría a ser un grupo aditivo. Es éste el sentido en que los conceptos de grupo aditivo y grupo conmutativo son idénticos. Existen casos en los que cualquier persona, simplemente por razones de hábito, vacilaría en hacer un tal cambio de notación. Por ejemplo, en el caso de **Z**, pocos aceptarían representar la suma de dos enteros como xy. En el caso del grupo de traslaciones del plano que mencionaremos en la próxima sección, cualquiera de las dos notaciones parece aceptable: la multiplicativa es más natural si pensamos en una traslación como una aplicación del plano en sí mismo y en la resultante de dos traslaciones como un producto de aplicaciones; por el contrario, la notación aditiva es preferible si pensamos en una traslación como definida por un vector, caso en el que la resultante o suma de dos traslaciones corresponde a la suma de los vectores correspondientes.

En virtud de lo precedente, adoptaremos la siguiente convención de terminología: designaremos por *grupo* a un grupo aditivo o multiplicativo indistintamente. Los grupos aditivos o multiplicativos que son conmutativos serán llamados colectivamente *grupos conmutativos*.

Ejercicios

1. Si el conjunto G posee uno, dos, tres, cuatro o cinco elementos, toda operación de multiplicación en relación con la cual G sea un grupo

multiplicativo es conmutativa. Si G posee seis elementos, entonces G se puede convertir en un grupo multiplicativo no conmutativo.

2. Formular y demostrar los análogos para los grupos multiplicativos de los ejercicios 2 y 3 de la página 51.

3. Un grupo multiplicativo donde todos los elementos x son *involutorios*, esto es $x = x^{-1}$, es conmutativo.

§ 3. SUBGRUPOS

Dos grupos, como dos conjuntos, pueden compararse mediante la relación de inclusión. Con ciertas excepciones, cuando consideramos dos grupos G y H (los cuales supondremos aditivos, por ejemplo) tales que H es un subconjunto de G, ambas operaciones de adición actúan del mismo modo, esto es, dados $x, y \in H$, cuando calculamos $x + y$ obtenemos el mismo resultado que si, considerando a x e y como elementos de G, hacemos el cálculo usando la adición de G. Esto es lo que sucede (para citar uno entre muchos ejemplos) en el caso de los grupos aditivos $\mathbf{Z}$ y $\mathbf{Q}$, donde $\mathbf{Z} \subset \mathbf{Q}$. En cambio, en el ejemplo 4 de la página 49, tenemos que $\mathbf{Z}/p \subset \mathbf{Z}$, pero el valor de la suma de dos elementos $x, y \in \mathbf{Z}/p$ depende de nuestra decisión de considerarlos ya sea como elementos del grupo aditivo $\mathbf{Z}/p$ o del grupo aditivo $\mathbf{Z}$.

Decimos que un subconjunto H de un grupo G es un *subgrupo* de G cuando H es un grupo en relación con la operación (de adición o de multiplicación) de G aplicada a los elementos de H; o sea, en el caso aditivo, cuando H es un grupo con la correspondencia $(x, y) \mapsto x + y$, donde $x, y \in H$ y $x + y$ se calcula en G. Se procede en forma análoga para el caso multiplicativo, con $(x, y) \mapsto xy$. Tenemos, entonces:

Proposición 1. Para que el subconjunto H del grupo aditivo G sea un subgrupo de G es necesario y suficiente que:

1. $0 \in H$,

2. $x, y \in H \Rightarrow x + y \in H$,

3. $x \in H \Rightarrow -x \in H$.

Demostración. Supongamos, en primer lugar, satisfechas las condiciones 1, 2 y 3. Por 2 vemos que la correspondencia $(x, y) \mapsto x + y$, donde $x, y \in H$ y $x + y$, se calcula en el sentido de G y actúa de $H \times H$ en H. Se satisface, pues, el axioma 1, página 47. Además $x + y = y + x$ para cualesquiera $x, y \in H$, pues las sumas se calculan en el sentido de G y la ley conmutativa vale entre dos elementos cualesquiera de G. Luego, se satisface el axioma 2, página 47. Idem para el axioma 3, página 47. Por la condición 1 de la proposición, tenemos que $0 \in H$. Como $x + 0 = x$ cualquiera que sea $x \in H$ (porque la suma se calcula en G, donde esta igualdad es verdadera), vemos que H tiene un elemento que cumple el papel de cero (a saber, el mismo cero de G); esto es, el axioma 4, página 47, también se satisface. Finalmente, si $x \in H$, entonces $-x \in H$ por la condición 3 del enunciado, y como $x + (-x) = 0$, se concluye que todo elemento de H posee un

simétrico en H (el cual, por cierto, es el propio simétrico de x en G). Por consiguiente, se satisface el axioma 5, página 48. Esto prueba que H es un grupo aditivo.

Recíprocamente, supongamos que H sea un grupo aditivo con relación a la correspondencia $(x, y) \mapsto x + y$, donde $x, y \in H$ y $x + y$ se calcula en el sentido de G. Por el axioma 1, página 47, esta correspondencia debe actuar de $H \times H$ en H, o sea que si $x, y \in H$, debemos tener $x + y \in H$, lo que prueba la condición 2 del enunciado. Además, H, como cualquier grupo aditivo, posee un elemento cero. Representemos ese elemento por $0'$ a fin de evitar *a priori* su confusión con el 0 de G. Entonces, $x + 0' = x$ para todo $x \in H$. Haciendo $x = 0'$, tenemos $0' + 0' = 0'$ y, como la suma se calcula en G, concluimos que $0' = 0$ (proposición 1, punto 3, página 49). Ahora bien, $0' \in H$. Luego, $0 \in H$, lo que establece la condición 1 del enunciado y, al mismo tiempo, muestra que no hay distinción entre los ceros del grupo y del subgrupo. Finalmente, a todo $x \in H$ le corresponde un elemento de H, que representaremos por $(-x)'$ para no confundirlo *a priori* con el simétrico $-x$ de x en G y tal que $x + (-x)' = 0'$, o sea, $x + (-x)' = 0$. Como la suma en cuestión es calculada en G, concluimos, por la unicidad del simétrico (pág. 48). que $(-x)' = -x$. Ahora bien, $(-x)' \in H$, lo que prueba que $-x \in H$, como lo requería la condición 3 del enunciado, además de mostrar que no hay razón para distinguir los simétricos del grupo y del subgrupo. QED

58 Es evidente que G es un subgrupo de sí mismo y que el conjunto constituido sólo por el cero de G es también un subgrupo de G. Así, pues, G es el mayor subgrupo de sí mismo y $\{0\}$ es el menor de los subgrupos de G.

Ejemplo 1. Cada uno de los tres primeros grupos aditivos **Z**, **Q**, **R** y **C** es un subgrupo del siguiente. Para mencionar otro grupo aditivo de números, recordemos que un número complejo x se dice *algebraico* cuando satisface, por lo menos, una ecuación algebraica:

$$x^m + a_1 x^{m-1} + \dots + a_m = 0$$

donde $a_1, \dots, a_m$ son números racionales y $m \geq 1$. Es claro que el número 0 es algebraico. Si

$$x^m + a_1 x^{m-1} + \dots + a_m = 0, \qquad y^n + b_1 y^{n-1} + \dots + b_n = 0$$

fueran dos ecuaciones algebraicas, con raíces $x_1, x_2, \dots, x_m$ e $y_1, y_2, \dots, y_n$, respectivamente, entonces, los números $x_i + y_j$ ($i = 1, \dots, m$; $j = 1, \dots, n$) son las raíces de una ecuación

$$z^{mn} + c_1 z^{mn-1} + \dots + c_{mn} = 0$$

y, como se demuestra en la teoría elemental de las ecuaciones, los coeficientes (c) se expresan como polinomios en los coeficientes (a) y (b): en particular, aquéllos serán números racionales si éstos lo fueran. Luego, la suma de dos números complejos algebraicos es también un número algebraico. Finalmente, es claro que el simétrico $-x$ de todo número complejo algebraico es también un número algebraico.

En resumen, el conjunto de los números complejos algebraicos es un subgrupo del grupo aditivo C de los números complejos.

Ejemplo 2. Consideremos el grupo aditivo F de las funciones reales definidas en $[a, b]$ (pág. 46). Afirmamos que el conjunto C de las funciones continuas en $[a, b]$ es un subgrupo de F. En efecto, la función idénticamente nula en $[a, b]$ es continua. Como se sabe, la suma $f + g$ de dos funciones continuas $f, g \in F$ también es continua. Finalmente, para toda $f \in F$ continua, la función $-f$ es también continua, lo que prueba nuestra aserción. Otro ejemplo de subgrupo de F es dado por el conjunto de las funciones reales derivables en $[a, b]$.

Para probar que un subconjunto es un subgrupo, a continuación se expone una forma más breve que la indicada en la proposición 1.

Proposición 2. Para que el subconjunto H del grupo aditivo G sea un subgrupo de G, es necesario y suficiente que:

1. $0 \in H$,

2. $x, y \in H \Rightarrow x - y \in H$.

Demostración. Supongamos satisfechas las condiciones 1, 2 y 3 de la proposición 1. Basta probar la condición 2 de arriba. Ahora bien, $x, y \in H$ implican $x, -y \in H$, donde $x + (-y) \in H$, o sea $x - y \in H$, como queríamos.

Recíprocamente, supongamos satisfechas las condiciones 1 y 2 del enunciado. Tenemos que establecer las condiciones 2 y 3 de la proposición 1. Si $x \in H$, como por hipótesis $0 \in H$, concluimos que $0 - x \in H$, o sea $-x \in H$, como requiere 3. Además, si $x, y \in H$, entonces, por lo que acabamos de probar, $x, -y \in H$, de donde $x - (-y) \in H$, o sea $x + y \in H$. QED

Hasta aquí hemos considerado la suma $x + y$ de apenas dos elementos de un grupo aditivo. Vamos ahora a extender la noción de suma a las partes de G, lo que no sólo permitirá una reformulación interesante de las proposiciones precedentes, sino que también encontrará su aplicación en el concepto de grupo cociente. Definiremos, pues, la suma $X + Y$ de dos subconjuntos $X, Y \subset G$ mediante

$$X + Y = \{x + y;\ x \in X,\ y \in Y\},$$

o sea, $X + Y$ es el conjunto de los elementos de G de la forma $x + y$, donde x varía en X e y recorre Y. Notemos que, en particular, uno de los dos conjuntos X e Y podría reducirse a un punto: por ejemplo, si X consistiera solamente del punto x, entonces $x + Y$ sería el conjunto de los elementos de la forma $x + y$, donde y varía en Y. Análogamente, definiremos $X - Y$ y $-X$ mediante

$$X - Y = \{x - y;\ x \in X,\ y \in Y\}, \qquad -X = \{-x;\ x \in X\}.$$

Con estas nuevas notaciones, las proposiciones 1 y 2 que preceden se convierten en:

Proposición 1. Para que el subconjunto H del grupo aditivo G sea un subgrupo de G es necesario y suficiente que:

1. $0 \in H$,
2. $H + H \subset H$,
3. $-H \subset H$.

Proposición 2. Para que el subconjunto H del grupo aditivo G sea un subgrupo de G es necesario y suficiente que:

1. $0 \in H$,
2. $H - H \subset H$.

En el caso de los grupos multiplicativos, las proposiciones análogas a 1 y 2 son las siguientes:

Proposición 3. Para que el subconjunto H del grupo multiplicativo G sea un subgrupo de G es necesario y suficiente que:

1. $e \in H$,
2. $x,\ y \in H \Rightarrow xy \in H$,
3. $x \in H \Rightarrow x^{-1} \in H$.

Proposición 4. Para que el subconjunto H del grupo multiplicativo G sea un subgrupo de G es necesario y suficiente que:

1. $e \in H$,
2. $x,\ y \in H \Rightarrow xy^{-1} \in H$.

Obsérvese que en esta proposición se puede sustituir la condición 2 por:

2. $x,\ y \in H \Rightarrow x^{-1}y \in H$.

Es claro también que G es un subgrupo de sí mismo y que el conjunto que se reduce a la unidad de G es también un subgrupo de G. Un subgrupo de un grupo conmutativo es también conmutativo.

Si se introducen las notaciones:

$$XY = \{xy;\ x \in X,\ y \in Y\},\quad X^{-1} = \{x^{-1};\ x \in X\},$$

donde $X,\ Y \in G$, las condiciones de la proposición 3 pasan a ser formuladas así:

1. $e \in H$,
2. $HH \subset H$,
3. $H^{-1} \subset H$.

y las de la proposición 4 se convierten en:

1. $e \in H$,

2. $HH^{-1} \subset H$.

Ejemplo 3. El grupo multiplicativo $\mathbf{Q}^*$ es un subgrupo de $\mathbf{R}^*$ y de $\mathbf{C}^*$. De modo análogo, $\mathbf{R}^*$ es un subgrupo de $\mathbf{C}^*$ (págs. 51-52). Los conjuntos $\mathbf{Q}_+^*$ y $\mathbf{R}_+^*$ de los números racionales > 0 y de los números reales > 0 son subgrupos de $\mathbf{Q}^*$ y $\mathbf{R}^*$.

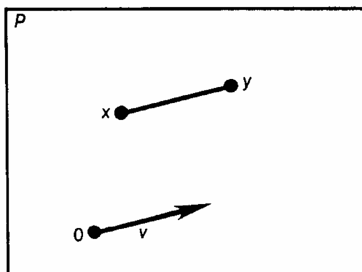


Fig. 31

Ejemplo 4. Sea P un plano euclideano. Consideremos un punto fijo O y el conjunto S_0 de los segmentos orientados de P con origen en O. Dado $v \in S_0$, designemos con t_v a la *traslación* de v en P, o sea, la aplicación $t_v: P \to P$ de P en sí mismo, que a cada punto $x \in P$ le asocia el punto $y \in P$ definido por la condición de que el segmento orientado $\overline{xy}$ sea equipolente a v. Toda traslación en P es una permutación del plano. El conjunto de las traslaciones en P es un subgrupo del grupo multiplicativo de las permutaciones de P (ejemplo 2, pág. 52). De hecho, la transformación idéntica $I: P \to P$ es una traslación en P, pues $I = t_0$, donde 0 es el elemento cero de S_0 (ejemplo 2, pág. 46). Además, se verifica fácilmente que

$$t_{v'}\, t_v = t_{v+v'},$$

lo que prueba que el producto de dos traslaciones es también una traslación. Finalmente, el inverso de una traslación es una traslación, pues

$$(t_v)^{-1} = t_{-v}.$$

Por ello, el conjunto de las traslaciones en P constituye el *grupo de las traslaciones en P*, el cual es evidentemente conmutativo. Conviene destacar que este grupo no depende de la elección del punto O.

Ejemplo 5. Consideremos un plano euclideano P y un punto O de P. Para todo ángulo α, indiquemos la rotación alrededor de O en un ángulo α por $r_{O,\alpha}$ o, simplemente, por r_α cuando O esté sobrentendido. Cada una de tales rotaciones es una permutación de P. Si se mantiene O fijo

y se hace variar α, el conjunto de las rotaciones alrededor de 0 constituye un subconjunto del grupo de las permutaciones de P. Esto se sigue de las relaciones siguientes:

$$I = r_0, \quad r_{\alpha'} r_\alpha = r_{\alpha+\alpha'}, \quad (r_\alpha)^{-1} = r_{-\alpha'},$$

donde $I: P \to P$ es la transformación idéntica.

Igualmente, si para todo punto O del plano y todo número real $k \neq 0$, se indicara con $h_{0,k}$ (o, simplemente, h_k) la homotecia de centro O y razón k, vemos que cada una de tales homotecias es una permutación de P y que el conjunto de las homotecias relativas a un mismo punto O constituye un subgrupo del grupo de las permutaciones de P, pues

$$I = h_1, \quad h_{k'} h_k = h_{kk'}, \quad (h_k)^{-1} = h_{1/k}.$$

Los grupos de las rotaciones y de las homotecias de centro común en un plano son obviamente conmutativos.

Ejemplo 6. Consideremos un plano euclideano P y un punto fijo O de P. En geometría elemental, se dice que dos figuras son semejantes cuando es posible pasar de una a otra por medio de una rotación alrededor de O, seguida de una homotecia de razón no nula y centro O y de una traslación. La semejanza de dos figuras no depende de la elección del punto O. Fijemos, pues, nuestra atención en las transformaciones del plano P en sí mismo, que pueden obtenerse por medio de una rotación alrededor de O, seguida de una homotecia de razón no nula y centro O y de una traslación. A una tal transformación de P se le da el nombre de *semejanza* en P:

$$\text{semejanza} = \text{traslación} \cdot \text{homotecia} \cdot \text{rotación}.$$

Una semejanza es una permutación de P, pues es el producto de tres permutaciones de P (proposición 4, pág. 22). ¿Por qué efectuamos primero una rotación, después una homotecia y al final una traslación? En realidad el orden no tiene mayor importancia. En efecto, es claro

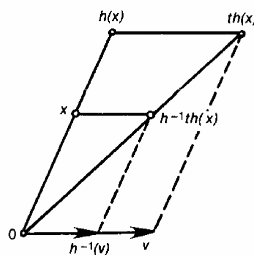


Fig. 32

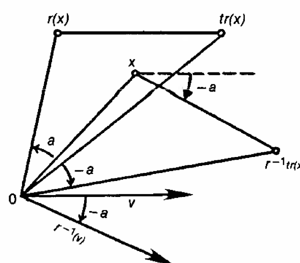


Fig. 33

que una rotación y una homotecia con el mismo centro siempre conmutan, de modo que el orden entre rotaciones y homotecias no tiene

ninguna importancia. Por otro lado, es fácil verificar con ejemplos que una traslación no necesariamente conmuta con una homotecia. Esta dificultad, empero, puede ser eludida del siguiente modo. Si t es una traslación y h es una homotecia, entonces $h^{-1}th$ es una traslación. En efecto, llamando v al segmento orientado de origen O que define a t y k a la razón no nula de la homotecia de centro O, entonces $h^{-1}th$ es la traslación definida por el segmento $h^{-1}(v)$, obtenido sometiendo a v a una homotecia de centro O y razón $1/k$, como se verifica fácilmente (Fig. 32). Análogamente, si r es una rotación, entonces $r^{-1}tr$ es una traslación, a saber: la traslación definida por el segmento $r^{-1}(v)$, obtenido sometiendo v a una rotación de $-\alpha$ alrededor del centro O de r, donde α es el ángulo que define a r (Fig. 33). Visto lo anterior, supongamos ahora que $s = thr$ es una semejanza, donde t es una traslación y h y r una homotecia de razón no nula y una rotación, ambas de centro O. Escribiendo $t' = h^{-1}th$, se obtiene una nueva traslación y, entonces, tenemos $ht' = th$, con $s = ht'r$, lo que prueba que es posible permutar la traslación con la homotecia, siempre que se sustituya la traslación dada con una nueva traslación. Con un raciocinio perfectamente similar, es posible permutar la traslación con la rotación. Por tanto, una semejanza puede considerarse como un producto de los tres tipos de transformaciones consideradas, tomadas en algún orden a nuestra elección. Afirmamos ahora que el conjunto S de las semejanzas es un subgrupo del de las permutaciones de P. En efecto, la transformación idéntica $I: P \to P$ es una semejanza, pues $I = I \cdot I \cdot I$ e I es tanto una traslación (de segmento nulo), como una homotecia (de razón uno) o una rotación (de ángulo nulo). Además, si $s = thr$ y $s' = t'h'r'$ son dos semejanzas (donde la notación obvia cualquier comentario), tendremos 63

$$s's^{-1} = t'h'r'r^{-1}h^{-1}t^{-1} = t'(h'h^{-1})(r'r^{-1})t^{-1},$$

pues la homotecia h^{-1} conmuta con las rotaciones r^{-1} y r'. Como $r'r^{-1}$ es una rotación y t^{-1} una traslación, existe una traslación t_1 tal que

$$(r'r^{-1})t^{-1} = t_1(r'r^{-1}),$$

donde

$$s's^{-1} = t'(h'h^{-1})t_1(r'r^{-1}).$$

De la misma forma, existe una traslación t_2 tal que

$$(h'h^{-1})t_1 = t_2(h'h^{-1}),$$

de donde se sigue

$$s's^{-1} = (t't_2)(h'h^{-1})(r'r^{-1})$$

y, por tanto, $s's^{-1}$ es una semejanza. En virtud de la proposición 4, página 60, se concluye que el conjunto de las semejanzas es un subgrupo del grupo de las permutaciones de P. El grupo de las semejanzas es un ejemplo importante, tomado de la geometría elemental, de un grupo no conmutativo. Como se dijo ya, la elección del punto O no influye en la naturaleza del grupo de las semejanzas. Este contiene,

además, los grupos de las traslaciones, de las rotaciones alrededor de un punto y de las homotecias de razón no nula relativas a un punto, como sus subgrupos, pues $t = t \cdot I \cdot I$, $h = I \cdot h \cdot I$ y $r = I \cdot I \cdot r$.

De un modo general, se da el nombre de *grupo de transformaciones* de un conjunto E a todo subgrupo del grupo de permutaciones de E. Los ejemplos 4, 5 y 6 constituyen, pues, grupos de transformaciones de un plano. La geometría ofrece un gran número de ejemplos importantes de grupos de transformaciones: el grupo de las afinidades, el grupo de las homografías, el grupo de las transformaciones conformes, etc.

Ejemplo 7. El conjunto $\mathbf{T}$ de los números complejos de módulo unidad es un subgrupo del grupo multiplicativo $\mathbf{C}^*$ de los números complejos no nulos (ejemplo 1, pág. 51). En efecto $|1| = 1$; y, si $|u| = 1$ y $|v| = 1$, entonces $|uv^{-1}| = |u| \cdot |v|^{-1} = 1$.

Del mismo modo, el conjunto de las raíces p-ésimas de la unidad es un subgrupo de $\mathbf{C}^*$ y también de $\mathbf{T}$. En efecto, $1^p = 1$; y si $u^p = v^p = 1$, entonces $(uv^{-1})^p = u^p(v^p)^{-1} = 1$.

Finalmente, el conjunto de las raíces complejas de la unidad (de orden no especificado) es un subgrupo de $\mathbf{C}^*$ y también de $\mathbf{T}$. En efecto, $1^1 = 1$ y, si $u^p = 1$, $v^q = 1$, entonces $(uv^{-1})^{pq} = (u^p)^q\,(v^q)^p = 1$.

Antes de concluir esta sección, conviene mencionar que un grupo no puede considerarse completamente conocido hasta que no se haya encontrado un modo de describir todos sus subgrupos, lo que no es fácil, aun en los casos corrientes. Entre los grupos cuyos subgrupos pueden describirse explícitamente de modo simple figura el grupo aditivo de los enteros $\mathbf{Z}$.

Proposición 5. Dado un entero natural p, el conjunto $\{np;\ n \in \mathbf{Z}\}$ de sus múltiplos enteros constituye un subgrupo del grupo aditivo $\mathbf{Z}$. Recíprocamente, a todo subgrupo H de $\mathbf{Z}$ le corresponde uno, y sólo un, entero natural p, tal que H es el conjunto de los múltiplos enteros de p.

Demostración. Dado el entero natural p, sea $H = \{np;\ n \in \mathbf{Z}\}$ el conjunto de los múltiplos enteros de p. Es claro que $0 \in H$, pues $0 = 0p$. Si x, $y \in H$, esto es, si $x = mp$ y $y = np$ $(m,\ n \in \mathbf{Z})$, entonces $x - y = (m - n)p$, luego $x - y \in H$, lo que prueba que H es un subgrupo de $\mathbf{Z}$ (proposición 2, pág. 60).

Recíprocamente, sea H un subgrupo cualquiera de $\mathbf{Z}$. Si H se redujera al número 0, basta con tomar $p = 0$, y es claro, entonces, que H es el conjunto de los múltiplos de p (y que no hay otra elección posible). Supongamos ahora que H contiene otros números además de cero. Entonces H contiene al menos un elemento mayor que cero: en efecto, de la hipótesis sobre H se sigue que H contiene al menos un elemento x distinto de cero. O $x > 0$ o bien $x < 0$, en cuyo caso $-x > 0$, lo que también prueba nuestra aserción, puesto que H es un subgrupo de $\mathbf{Z}$ y $x \in H$ implica que $-x \in H$. Una de las propiedades fundamentales del conjunto de los enteros naturales es, como se sabe, la de que toda colección no vacía de enteros naturales contiene un elemento

menor que los demás en esa colección. Aplicando esta observación al conjunto de los elementos de H que son > 0, designemos por p el menor entero > 0 en H. Por definición, $p > 0$. Como $p \in H$ y H es un subgrupo de **Z**, vemos que $2p = p + p \in H$, $3p = p + 2p \in H$, etc. Además, $p \in H$ implica que $-p \in H$. Luego, $(-2)p = (-p) + (-p) \in H$, $(-3)p = (-p) + (-2p) \in$ $\in H$, etc. En resumen, $np \in H$ para cualquier $n \in \mathbf{Z}$. Además, todo elemento $x \in H$ es un múltiplo de p. Para comprobarlo, dividamos x entre p. Tendremos que $x = qp + r$, donde q y r son el cociente y el residuo. De ahí que $r = x - qp$ pertenezca a H, dado que x, $qp \in H$. Ahora, $0 \leq r < p$, pues r es el residuo de una división entre p. Esto, unido a $r \in H$ y a la definición de p, exige $r = 0$, donde $x = qp$, como queríamos. Luego, H es el conjunto de los múltiplos de p.

Supongamos ahora que p y p' sean dos enteros naturales tales que H sea tanto el conjunto de los múltiplos de p como el conjunto de los múltiplos de p'. Como p es múltiplo de sí mismo, vemos que $p \in H$ y, por tanto, p es múltiplo de p'. Igualmente, p' es múltiplo de p. Conclusión: $p = p'$. QED

El entero natural p que determina (y es determinado) por todo subgrupo H de **Z**, de acuerdo con la proposición anterior, se denomina *generador natural* de H.

Ejercicios

1. Sean G, H e I tres grupos tales que $I \subset H \subset G$. Si I es un subgrupo de G y H es un subgrupo de G, entonces I es un subgrupo de H.

2. En las proposiciones 1 y 2, §3, se puede sustituir la condición de que $0 \in H$ por la condición de que H no sea vacío. También se puede sustituir la condición $H + H \subset H$ por $H + H = H$, $-H \subset H$ por $-H = H$ y $H - H \subset H$ por $H - H = H$. Lo mismo para las proposiciones 3 y 4.

3. Sea G un grupo aditivo. Si $X, Y, \ldots \subset G$, mostrar que

   1. $X + Y = Y + X$,

   2. $X + (Y + Z) = (X + Y) + Z$,

   3. $X \subset X'$, $Y \subset Y' \Rightarrow X + Y \subset X' + Y'$.

Análogamente para el caso multiplicativo (limitando 1. al caso conmutativo).

4. El conjunto $\mathcal{P}(G)$ de las partes de G no es un grupo con relación a la operación (de adición o multiplicación) extendida a las partes.

5. La intersección de un número finito de subgrupos de **Z** diferentes de 0 es también diferente de 0.

6. Dado el entero $p \geq 2$, si $n \geq 1$ fuera un divisor de p (digamos, $p = qn$), entonces el conjunto constituido por 0, n, $2n, \ldots, (q-1)n$ es un subgrupo del grupo aditivo $\mathbf{Z}/p$ y, de esta forma, se establece una correspondencia biunívoca entre los divisores $n \geq 1$ de p y los subgrupos del grupo aditivo $\mathbf{Z}/p$ diferentes de 0.

7. Consideremos un plano euclideano P y un punto 0 de P. Toda semejanza s en P puede ser escrita de un modo único como un producto $s = thr$ de una traslación t por una homotecia h de centro 0 y razón > 0, y una rotación r alrededor de 0. Idem para las descomposiciones en otros órdenes, como $s = htr$, etc.

8. Sea A un conjunto no vacío de enteros ≥ 1 con la propiedad siguiente: $m, n \in A$ implica $mn \in A$ (un ejemplo usual se obtiene considerando un entero p y tomando el conjunto A formado por 1, p, p^2, $p^3, \ldots$). El conjunto de los números racionales de la forma m/n, donde $m \in \mathbf{Z}$ y $n \in A$, es un subgrupo del grupo aditivo $\mathbf{Q}$.

9. Sea A un conjunto de enteros ≥ 1 con las siguientes propiedades: 1) $1 \in A$, 2) si $m, n \in A$, entonces m. c. m. $(m, n) \in A$ (donde m. c. m. denota el mínimo común múltiplo); 3) si $m \geq 1$ divide a $n \in A$, entonces $m \in A$. Entonces el conjunto de los números complejos z, tales que $z^m = 1$ para al menos un $m \in A$ es un subgrupo del grupo multiplicativo de las raíces complejas de la unidad. Recíprocamente, todo subgrupo de este grupo corresponde a un conjunto A, y sólo a uno, del tipo indicado.

§ 4. HOMOMORFISMOS E ISOMORFISMOS

Cuando se estudian los conjuntos, las transformaciones que se consideran entre los varios conjuntos no están sujetas a restricciones.

Cuando, empero, los dominios y contradominios de las funciones son conjuntos en los cuales están dadas ciertas operaciones algebraicas, pueden destacarse de entre esas funciones a aquellas que *respetan* esas operaciones. Así, somos llevados a la noción de homomorfismo --o isomorfismo, en el caso de la biunivocidad-- que se definirá entre grupos y, más adelante, se repetirá entre otros sistemas algebraicos.

Sean G y H dos grupos, inicialmente supuestos aditivos. Un *homomorfismo* de G en H es una función $f: G \to H$, definida en G y con valores en H, tal que

$$f(x + y) = f(x) + f(y), \quad (x, y \in G),$$

o, como es costumbre decir, tal que la transformación de la suma sea la suma de las transformaciones. Esta condición, que define a los homomorfismos, se denomina propiedad *aditiva*. Notemos que

$$f(0) = 0. \quad f(-x) = -f(x)$$

para todo homomorfismo f. En efecto, escribiendo $x = 0$ e $y = 0$ en la ecuación $f(x + y) = f(x) + f(y)$, tenemos $f(0) = f(0) + f(0)$, donde $f(0) = = 0$, por la proposición 1, página 49. Además, $f(x) + f(-x) = f[x + + (-x)] = f(0) = 0$, lo que prueba $f(-x) = -f(x)$ (pág. 48). Notemos también que

$$f(x - y) = f(x) - f(y),$$

pues $f(x - y) = f[x + (-y)] = f(x) + f(-y) = f(x) + [-f(y)] = f(x) - f(y)$.

Por definición, un *isomorfismo* de G en H es todo homomorfismo inyectivo de G en H. Por analogía con la terminología adoptada en el caso de conjuntos (pág. 16), cuando el isomorfismo $f: G \to H$ fuera *sobre*, diremos que f es un isomorfismo *entre* G *y* H o *de* G *sobre* H. La existencia de al menos un isomorfismo $f: G \to H$ entre G y H significa que, desde el punto de vista algebraico, los grupos G y H se comportan del mismo modo. Los dos grupos G y H se dicen *isomorfos* cuando existe al menos un isomorfismo de G sobre H. Se acostumbra decir, entonces, que G y H *son iguales salvo por isomorfismos*.

De una manera más general, se dice que el grupo H es una imagen *homomorfa* de G cuando existe al menos un homomorfismo de G sobre H.

Nótese que, dados G y H, la función $0: G \to H$ (pág. 13), que a cada elemento de G asocia siempre el cero de H, es un homomorfismo llamado homomorfismo *cero* de G en H.

Ejemplo 1. Consideremos el grupo aditivo $\mathbf{R}$ de los números reales. Dado $a \in \mathbf{R}$, la función de $\mathbf{R}$ en $\mathbf{R}$ definida mediante $x \mapsto ax$ es un homomorfismo de $\mathbf{R}$ en $\mathbf{R}$, pues $f(x + y) = a(x + y) = ax + ay = f(x) + f(y)$. Si $a = 0$, es claro que f es el homomorfismo 0. Si $a \neq 0$, entonces f es un isomorfismo de $\mathbf{R}$ sobre sí mismo. Además, hay un teorema clásico de Cauchy, que afirma que toda *función* $f: \mathbf{R} \to \mathbf{R}$ continua, tal que $f(x + y) = f(x) + f(y)$ para cualesquiera $x, y \in \mathbf{R}$, es de la forma $f(x) = ax$. En otros términos, todo homomorfismo continuo de $\mathbf{R}$ en sí mismo es del tipo $x \mapsto ax$.

Ejemplo 2. En el grupo aditivo $\mathbf{C}$ de los números complejos, la correspondencia $z \to \overline{z}$ que a cada z asocia su conjugado $\overline{z}$ es un isomorfismo de $\mathbf{C}$ sobre sí mismo. En efecto, como es sabido,

$$\overline{z + w} = \overline{z} + \overline{w}, \qquad (z, w \in \mathbf{C})$$

lo que prueba que la función $z \mapsto \overline{z}$ es un homomorfismo. Como $\overline{\overline{z}} = z$, es fácil concluir que $z \mapsto \overline{z}$ es una permutación (además, involutoria) de $\mathbf{C}$. Luego, $z \mapsto \overline{z}$ es un isomorfismo de $\mathbf{C}$ sobre sí mismo.

Ejemplo 3. Consideremos el conjunto C de las funciones reales continuas en el intervalo cerrado acotado $[a, b]$. Recordemos (ej. 2, pág. 59) que C es un subgrupo del grupo aditivo F de las funciones reales en $[a, b]$. Por tanto, C es un grupo aditivo con relación a la adición usual de funciones reales. La correspondencia de C en $\mathbf{R}$ dada por

$$f \mapsto \int_a^b f(x)\, dx$$

que a cada $f \in C$ le asocia su integral en $[a, b]$, es un homomorfismo de C en $\mathbf{R}$, pues

$$\int_a^b [f(x) + g(x)]\, dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx;$$

este homomorfismo no es un isomorfismo.

Ejemplo 4. Como se advirtió ya (pág. 59), el conjunto D de las funciones reales derivables en $[a, b]$ constituye un subgrupo del grupo aditivo $\mathbf{F}$ de las funciones reales en $[a, b]$. En consecuencia, D es un grupo aditivo con la suma usual de funciones reales. La correspondencia de D en $\mathbf{F}$, dada por

$$f \mapsto df/dx$$

que a cada $f \in D$ le asocia su derivada, es un homomorfismo de D en $\mathbf{F}$, pues

$$\frac{d(f + g)}{dx} = \frac{df}{dx} + \frac{dg}{dx}$$

Este homomorfismo no es un isomorfismo.

Ejemplo 5. Consideremos los grupos aditivos $\mathbf{Z}$ y $\mathbf{Z}/p$, el primero respecto a la adición usual y el segundo respecto a la adición módulo p, para un $p \geq 1$ dado. La función $r: \mathbf{Z} \to \mathbf{Z}/p$ que a cada $x \in \mathbf{Z}$ le asocia el residuo de dividir x entre p, es un homomorfismo, esto es,

$$r(x + y) = r(x) \overset{p}{+} r(y), \quad (x, y \in \mathbf{Z}).$$

En efecto, dividiendo x e y entre p, tendremos

$$x = qp + m, \quad y = q'p + m',$$

donde q y q' son los cocientes y m y m' los residuos. Entonces, $r(x) = m$ y $r(y) = m'$. Para calcular $m \overset{p}{+} m'$, se divide $m + m'$ entre p, lo que da

$$m + m' = q''p + m'',$$

donde q'' es el cociente y m'' es el residuo. De ahí, $m'' = m \overset{p}{+} m'$. Ahora,

$$x + y = (q + q')p + (m + m') = (q + q' + q'')p + m''$$

y, en virtud de la definición de residuo de una división, se concluye que

$$r(x + y) = m'' = m \overset{p}{+} m' = r(x) \overset{p}{+} r(y),$$

como se deseaba.

Los ejemplos dados arriba ilustran la variedad de casos en matemática elemental en los que aparece la noción de homomorfismo. El último es el único que no se acostumbra a presentar como posible candidato a ejemplo de homomorfismo.

En el caso de grupos multiplicativos G y H, un *homomorfismo* es una función $f: G \to H$, que posee la siguiente propiedad *multiplicativa*

$$f(xy) = f(x)\,f(y), \quad (x, y \in G).$$

Notemos que, entonces

$$f(e) = e, \qquad f(x^{-1}) = [f(x)]^{-1}.$$

La demostración es análoga a la del caso aditivo y, por tanto, será omitida. La función que a todo elemento de G le asocia la unidad de H es un homomorfismo que por eso se denomina *homomorfismo unidad*.

Además de los casos en que tanto G como H son grupos aditivos o multiplicativos, podemos considerar dos casos mixtos: el de G aditivo y H multiplicativo y el de G multiplicativo y H aditivo. En el primer caso, un homomorfismo $f: G \to H$ es una función tal que

$$f(x + y) = f(x)f(y), \qquad (x, y \in G);$$

y, en el segundo, una función $f: G \to H$ para la cual

$$f(xy) = f(x) + f(y), \qquad (x, y \in G).$$

En cualquier caso, se define un isomorfismo de G en H como un homomorfismo biunívoco de G en H. A todo isomorfismo de G sobre H se le da el nombre de isomorfismo entre G y H. Cuando existe al menos un isomorfismo de G sobre H, estos dos grupos se dicen *isomorfos* o *iguales salvo por isomorfismos*. De una manera general, se dice que H es una imagen homomorfa de G si existe, al menos, un homomorfismo de G sobre H.

Es claro que

$$f(0) = e, \ f(-x) = f(x)^{-1} \text{ o } f(e) = 0, \ f(x^{-1}) = -f(x)$$

conforme al caso considerado. También se emplean las denominaciones de homomorfismo *cero* o *unidad*, según sea H aditivo o multiplicativo.

Ejemplo 6. Consideremos un plano euclideano P y un punto O de P. A cada número real $k \neq 0$, asociemos una homotecia h_k de P de centro O y razón k. Sabemos que

$$h_{kk'} = h_k h_{k'}$$

--o $h(kk') = h(k)h(k')$, si hubiésemos empleado la notación $h(k)$ en vez de h_k-- o sea, la correspondencia $k \mapsto h_k$ es un homomorfismo del grupo multiplicativo $\mathbf{C}^*$ (pág. 52) en el grupo multiplicativo de las homotecias de centro O (pág. 62). Es claro que esta correspondencia es también un isomorfismo entre los dos grupos.

Ejemplo 7. Indiquemos con $\mathbf{R}_+^*$ el grupo multiplicativo de los números reales > 0 (pág. 61) y sea $a \in \mathbf{R}$ un número real fijo. La función $f: \mathbf{R}_+^* \to \mathbf{R}_+^*$ dada por $x \mapsto x^a$ es un isomorfismo de $\mathbf{R}_+^*$ sobre sí mismo, pues $f(xy) = (xy)^a = x^a y^a = f(x)f(y)$. Cuando $a = 0$, f es el homomorfismo unidad. Si $a \neq 0$, entonces f es un isomorfismo de $\mathbf{R}_+^*$ sobre sí mismo.

Ejemplo 8. La correspondencia $z \mapsto \bar{z}$, que a todo número complejo $z \neq 0$ le asocia su conjugado, es un homomorfismo del grupo multiplicativo $\mathbf{C}^*$ (pág. 52) sobre sí mismo. Es lo que traduce la propiedad conocida.

$$\overline{zw} = \bar{z}\bar{w} \qquad (z, w \in \mathbf{C}^*).$$

Luego, de $\bar{\bar{z}} = z$ resulta que $z \mapsto \bar{z}$ es un isomorfismo de $\mathbf{C}^*$ sobre sí mismo.

Ejemplo 9. La correspondencia $x \mapsto |x|$ que a todo número real $x \neq 0$ le asocia su valor absoluto, es un homomorfismo del grupo multiplicativo $\mathbf{R}^*$ en el grupo multiplicativo $\mathbf{R}^*_+$, puesto que

$$|xy| = |x| \cdot |y| \qquad (x, y \in \mathbf{R}^*).$$

Análogamente, la correspondencia de $z \mapsto |z|$ que a todo número complejo z le asocia su módulo es un homomorfismo del grupo multiplicativo $\mathbf{C}^*$ en el grupo multiplicativo $\mathbf{R}^*_+$, pues

$$|zw| = |z| \cdot |w| \qquad (z, w \in \mathbf{C}^*).$$

Ambos actúan sobre $\mathbf{C}^*_+$, pero no son isomorfismos.

Ejemplo 10. Consideremos el grupo multiplicativo S de las seme-

janzas en un plano P (págs. 61-63). Dada la semejanza $s \in S$, escribamos $s = thr$, donde t es una traslación, h es una homotecia de razón $k \neq 0$ y r es una rotación. Representemos con $d(x, y)$ la distancia de dos puntos $x, y \in P$. Es claro que una traslación y una rotación no alteran las distancias, esto es

$$d[t(x), t(y)] = d(x, y) \text{ y } d[r(x), r(y)] = d(x, y), \qquad (x, y \in P).$$

Por otro lado, toda homotecia altera las distancias en una razón constante > 0, pues

$$d[h(x), h(y)] = |k| \cdot d(x, y), \qquad (x, y \in P).$$

De ahí resulta

$$d[(hr)(x), (hr)(y)] = d[h\{r(x)\}, h\{r(y)\}] = |k| \cdot d[r(x), r(y)] = \\ = |k| \cdot d(x, y)$$

y, por tanto,

$$d[(thr)(x), (thr)(y)] = d[t\{(hr)(x)\}, t\{(hr)(y)\}] = \\ = d[(hr)(x), (hr)(y)] = |k| \cdot d(x, y),$$

o sea,

$$d[s(x), s(y)] = |k| \cdot d(x, y), \qquad (x, y \in P),$$

lo que muestra que toda semejanza altera las distancias en una razón constante > 0. Esta razón constante, que obviamente está perfectamente determinada por s, se denomina *razón* de la semejanza s y se representa por $|s|$, y está caracterizada por la igualdad

$$d[s(x),\ s(y)] = |s| \cdot d(x,\ y), \qquad (x,\ y \in P).$$

La correspondencia $s \mapsto |s|$, que a cada semejanza asocia su razón, es una función de S en $\mathbf{R}_+^*$ y, más que eso, es un homomorfismo. En efecto

$$d[(ss')(x),\ (ss')(y)] = d[s\{s'(x)\},\ s\{s'(y)\}] = |s| \cdot$$
$$\cdot\ d[s'(x),\ s'(y)] = |s| \cdot |s'| \cdot d(x,\ y),$$

lo que prueba que $|ss'| = |s| \cdot |s'|$ $(s,\ s' \in S)$.

Ejemplo 11. La función $f\colon \mathbf{R} \to \mathbf{R}_+^*$ dada por $x \mapsto a^x$ (donde a es un número real > 0, pero $\neq 1$) es un isomorfismo del grupo aditivo $\mathbf{R}$ sobre el grupo multiplicativo $\mathbf{R}_+^*$, pues $f(x + y) = a^{x+y} = a^x a^y = f(x)f(y)$ y f es biunívoca de $\mathbf{R}$ sobre $\mathbf{R}_+^*$. La correspondencia inversa $g\colon \mathbf{R}_+^* \to \mathbf{R}$, dada por $x \mapsto \log_a x$, es un isomorfismo del grupo multiplicativo $\mathbf{R}_+^*$ sobre el grupo aditivo $\mathbf{R}$, pues $g(xy) = \log_a(xy) = \log_a x + \log_a y = g(x) + g(y)$ y g es biunívoca de $\mathbf{R}_+^*$ sobre $\mathbf{R}$. El grupo aditivo $\mathbf{R}$ y el grupo multiplicativo $\mathbf{R}_+^*$ son, pues, isomorfos.

Ejemplo 12. La correspondencia $z \mapsto a^z$ (donde $a > 0$ es un número real fijo) es un isomorfismo del grupo aditivo $\mathbf{C}$ sobre el grupo multiplicativo $\mathbf{C}^*$. Análogamente, la correspondencia $x \mapsto a^{ix}$ es un homomorfismo del grupo aditivo $\mathbf{R}$ sobre el grupo multiplicativo $\mathbf{T}$ (pág. 64). Por último, la correspondencia $x \mapsto a^{2\pi ix}$ es un homomorfismo del grupo aditivo $\mathbf{Q}$ sobre el grupo multiplicativo de las raíces complejas de la unidad (pág. 64).

Ejemplo 13. La correspondencia $a \to r_a$ que asocia a todo número real la rotación r_a alrededor del punto fijo O es un homomorfismo del grupo aditivo $\mathbf{R}$ sobre el grupo multiplicativo de las rotaciones alrededor de O (págs. 61-62).

A todo homomorfismo $f\colon G \to G$ de un grupo G en sí mismo, se le denomina *endomorfismo* de G. Llámase *automorfismo* de G a todo isomorfismo de G sobre sí mismo. Los casos 1 (pag. 67) y 7 (pág. 69) antes vistos constituyen ejemplos de endomorfismos y de automorfismos, si $a \neq 0$.

Obsérvese que la transformación idéntica $I\colon G \to G$ es siempre un automorfismo del grupo G. En todo grupo aditivo G, la correspondencia $x \mapsto -x$ es un automorfismo de G, pues $-(x + y) = (-x) + (-y)$ y esta correspondencia es biunívoca de G sobre sí mismo. Lo mismo puede decirse de la correspondencia $x \mapsto x^{-1}$ en el caso de un grupo multiplicativo conmutativo G. En la ausencia de conmutatividad, $f\colon x \to x^{-1}$ no es un homomorfismo de G en sí mismo, pues la igualdad que f satisface es $f(xy) = f(y)f(x)$ y no $f(xy) = f(x)f(y)$. De ahí la siguiente definición: Dados dos grupos multiplicativos G y H, llámase *antihomomorfismo* de G en H a toda función $f\colon G \to H$ tal que

$$f(xy) = f(y)f(x) \qquad (x,\ y \in G).$$

Las nociones de antiisomorfismo, antiendomorfismo y antiautomorfismo se definen de manera similar. Así, pues, $x \mapsto x^{-1}$ es un antiautomorfismo de todo grupo multiplicativo.

Proposición 1. Si $f: G \to H$ y $g: H \to I$ son homomorfismos entre los grupos G, H e I, entonces $gf: G \to I$ es también un homomorfismo.

Demostración. Para fijar la notación, supongamos que G, H e I son aditivos. Entonces

$$(gf)(x+y) = g\{f(x+y)\} = g\{f(x)+f(y)\} = g\{f(x)\} + g\{f(y)\} =$$
$$= (gf)(x) + (gf)(y)$$

para x, $y \in G$, como se quería.

Proposición 2. Si $f: G \to H$ es un isomorfismo del grupo G sobre el grupo H, entonces $f^{-1}: H \to G$ es un isomorfismo de H sobre G.

Demostración. Se sabe ya (pág. 16) que f^{-1} es biunívoca de H sobre G. Resta probar que f^{-1} es un homomorfismo. Para fijar las ideas vamos a suponer que G y H son aditivos. Dados u, $v \in H$, escribamos $x = f^{-1}(u)$, $y = f^{-1}(v)$. Entonces,

$$f(x+y) = f(x) + f(y) = u + v,$$

lo que prueba que $f^{-1}(u+v) = x + y$, o sea, $f^{-1}(u+v) = f^{-1}(u) + f^{-1}(v)$, como se quería. QED

Ejercicios

1. Todo subgrupo diferente de 0 del grupo aditivo $\mathbf{Z}$ es isomorfo a $\mathbf{Z}$.

2. Para que un homomorfismo $f: G \to H$, en el caso de grupos aditivos, sea un isomorfismo, es necesario y suficiente que $f(x) = 0 \Rightarrow x = 0$. De igual modo para el caso multiplicativo.

3. Todo endomorfismo del grupo aditivo $\mathbf{Z}$ es de la forma $x \mapsto a.c$, donde $a \in \mathbf{Z}$. En particular, todo endomorfismo diferente del endomorfismo 0 de $\mathbf{Z}$ es biunívoco. $\mathbf{Z}$ tiene apenas dos automorfismos, a saber: $x \mapsto x$ y $x \mapsto -x$.

4. Todo endomorfismo del grupo aditivo $\mathbf{Q}$ es de la forma $x \mapsto ax$, donde $a \in \mathbf{Q}$. En particular, todo endomorfismo diferente del endomorfismo 0 de $\mathbf{Q}$ es un automorfismo.

5. La correspondencia dada por

$$x \mapsto \exp(2\pi i x/p)$$

es un isomorfismo del grupo aditivo $\mathbf{Z}/p$ sobre el grupo multiplicativo de las raíces complejas p-ésimas de la unidad.

6. El grupo multiplicativo de las raíces complejas p-ésimas de la unidad posee p endomorfismos, a saber: $x \mapsto x^n (n = 0, 1, \ldots, p-1)$. El endomorfismo $x \mapsto x^n$ es un automorfismo si, y sólo si, n es primo relativo con p. En particular, si p es primo, y solamente en este caso, todo endomorfismo diferente del endomorfismo unidad es un automorfismo.

7. El conjunto de las funciones $x \mapsto az + b$ $(a, b, z \in \mathbf{C}, a \neq 0)$ es un subgrupo del grupo $\mathbf{C}!$ de las permutaciones de $\mathbf{C}$. El grupo constituido por estas funciones es isomorfo al grupo de las semejanzas de un plano euclideano.

8. Sea $f: G \to H$ un homomorfismo del grupo G en el grupo H. Para todo subgrupo $X \subset G$, $f(X)$ es un subgrupo de H. En particular, $f(G)$ es un subgrupo de H. Para todo subgrupo $Y \subset H$, $f^{-1}(Y)$ es un subgrupo de G. Si G es conmutativo, el grupo $f(G)$ es también conmutativo.

9. El conjunto de los automorfismos de un grupo G es un subgrupo del grupo $G!$ de las permutaciones de G. Lo mismo puede decirse del conjunto formado por los automorfismos y antiautomorfismos de G.

10. Sea G un grupo multiplicativo. Todo elemento $a \in G$ determina una permutación f_a: $G \to G$ de G, a saber: la permutación dada por $x \mapsto ax$. Demostrar que la correspondencia que a cada $a \in G$ asocia $f_a \in G!$ es un isomorfismo de G en $G!$ En particular, todo grupo multiplicativo es isomorfo a un grupo de transformaciones de un conjunto *(teorema de Cayley)*. La correspondencia que a cada $a \in G$ le asocia la permutación $x \mapsto xa$ del conjunto G es un antiisomorfismo de G en $G!$

11. Sea G un grupo aditivo. Consideremos también un conjunto H y una correspondencia de $H \times H$ en H que, a todo par (x, y), asocia el elemento $x + y$. Si $f: G \to H$ es una función tal que $f(x + y) = f(x) + f(y)$, entonces el conjunto $f(G)$ constituirá un grupo aditivo (aunque H no sea un grupo aditivo) en relación con la referida correspondencia. Idem para los grupos multiplicativos.

§ 5. MULTIPLOS Y POTENCIAS

Consideremos un grupo aditivo G. Dado $x \in G$, las expresiones $x + x$, $x + x + x, \ldots$ se representarán en forma abreviada por $2x$, $3x, \ldots$. De este modo, queda definido un producto nx de cada entero $n \geq 2$ por cada elemento $x \in G$. Se acostumbra extender la definición del producto nx a los valores restantes del entero n del siguiente modo. Para $n = 1$ y $n = 0$, se define $1x = x$ y $0x = 0$ (el 0 del segundo miembro designa al 0 de G). Para $n < 0$, se define $nx = -[(-n)x]$ y se observa que $(-n)x$ tiene sentido (porque $n < 0 \Rightarrow -n > 0$, que es un caso definido ya). Asociada, pues, a todo grupo aditivo G, se tiene una correspondencia $(n, x) \mapsto nx$ definida de $\mathbf{Z} \times G$ en G, que a cada entero n y cada x de G asocia el *múltiplo entero* $nx \in G$ de x.

Proposición 1. La correspondencia $(n, x) \mapsto nx$ de $\mathbf{Z} \times G$ en G posee las siguientes propiedades:

1. $(m + n)x = mx + nx$,

2. $m(x + y) = mx + my$,

3. $m(nx) = (mn)\, x$,

4. $1x = x$.

Demostración. Si $m \geq 2$ y $n \geq 2$, el punto 1 resulta inmediatamente de la definición de múltiplo. Si se examinan por separado los casos $m = 0$, 1 y $n = 0$, 1, es fácil concluir la validez de 1, para $m \geq 0$ y $n \geq 0$. Supongamos ahora que $m \geq 0$ y $n < 0$. Podemos tener $m + n \geq 0$ o $m + n <$ < 0. Para fijar ideas, supongamos $m + n \geq 0$. Entonces,

$$(m + n)x + (-n)x = [(m + n) + (-n)]x = mx,$$

donde $(m + n)x = mx + nx$, como queríamos. De forma semejante se completa la demostración de los casos restantes. La demostración de los puntos 2 y 3 se deja a cargo del lector. El punto 4 figura en la propia definición de múltiplo. QED

Más adelante, cuando se estudie el concepto de módulo sobre un anillo, se verá que la proposición anterior expresa que todo grupo aditivo es un módulo sobre $\mathbf{Z}$.

Otras propiedades, que se infieren fácilmente de las ya indicadas, son las siguientes:

$$0x = 0,\ m0 = 0,\ (-1)x = -x,\ (m - n)x = mx - nx,$$

$$m(x - y) = mx - my,\ (-m)x = m(-x) = -mx,\ (-m)(-x) = mx.$$

La noción pertinente al caso de un grupo multiplicativo G es la de la *potencia entera*. Si $x \in G$, las expresiones xx, xxx, ... se representarán en forma abreviada por x^2, x^3, ..., etc. La potencia x^n pasa, entonces, a tener sentido para cada entero $n \geq 2$ y cada elemento $x \in G$. La definición se completa para los demás valores de n, escribiendo $x^1 = x$ y $x^0 = e$ y $x^n = (x^{-n})^{-1}$ para $n < 0$, donde el exponente -1 representa el inverso. De este modo queda introducida una correspondencia $(n, x) \mapsto x^n$ de $\mathbf{Z} \times G$ en G, que a cada entero n y a cada $x \in G$ le asocia la *potencia entera* $x^n \in G$ de x.

Proposición 2. La correspondencia $(n, x) \mapsto x^n$ de $\mathbf{Z} \times G$ en G posee las siguientes propiedades:

1. $x^{m+n} = x^m x^n$,

2. $(xy)^m = x^m y^m$, si es que $xy = yx$,

3. $(x^n)^m = x^{mn}$,

4. $x^1 = x$.

La demostración es análoga a la de la proposición precedente y será omitida. Adviértase solamente que el punto 2 incluye la condición de que los elementos x e y conmuten. En particular, el punto 2 es válido para cualquier par de elementos de un grupo conmutativo. Se deja al lector la tarea de enunciar las propiedades de las potencias que son

análogas a las arriba indicadas para los múltiplos. Notemos que x^{-1} podría tener dos sentidos: la potencia de exponente -1 y el inverso multiplicativo. Ambas interpretaciones coinciden.

Las nociones de múltiplo o potencia entera dan lugar al concepto de orden de un elemento de un grupo. Como en matemática elemental se acostumbra mencionar este concepto explícitamente sólo en su versión multiplicativa, comenzaremos introduciéndolo en el caso de los grupos multiplicativos.

Sea G, pues, un grupo multiplicativo. Dado un elemento $x \in G$, examinemos el conjunto H de los enteros n tales que $x^n = e$, esto es, el conjunto de los exponentes n de las potencias de x iguales a la unidad. Afirmamos que este conjunto H de enteros es un subgrupo del grupo aditivo **Z**. En efecto, $x^0 = e$, esto es, $0 \in H$. Además, si $x^m = e$ y $x^n = e$, entonces $x^{m+n} = x^m x^n = ee = e$, esto es, $m, n \in H \Rightarrow m + n \in H$. Finalmente, si $x^m = e$, entonces $x^{-m} = (x^m)^{-1} = e^{-1} = e$, esto es, $m \in H \Rightarrow$ $\Rightarrow -m \in H$, lo que completa la demostración de nuestra aserción. Aplicando la proposición 5, página 64, vemos que existe un entero natural p, determinado por la propiedad de que H coincide con el conjunto de los múltiplos enteros de p. Dicho p recibe el nombre de *orden* de x. Resumiendo, el orden p de x es el entero natural caracterizado por las dos propiedades siguientes:

1. $x^p = e$ y, de un modo más general, $x^m = e$, para cualquier múltiplo $m \in \mathbf{Z}$ de p.

2. si $x^m = e$, $m \in \mathbf{Z}$, entonces m es un múltiplo de p.

Otra forma usual de enunciar estas dos propiedades es como sigue: dos potencias x^m y x^n (m, $n \in \mathbf{Z}$) coinciden si, y sólo si, $m \equiv n$ (mód. p). En efecto, $x^m = x^n$ equivale a $x^{m-n} = e$, que, a su vez, equivale a que $m - n$ sea un múltiplo de p.

El elemento x se dice *libre* cuando su orden p es igual a 0. En otros términos, si $x^m = e$ exige que $m = 0$; o, también, si $x^m = x^n$ sólo cuando $m = n$. Cuando el orden p es mayor que cero, x se dice *periódico* y, entonces, también se acostumbra llamar a p el *período* de x. En este caso, p es el menor entero > 0 tal que $x^p = e$. El elemento unidad e es periódico de período 1 y es también el único elemento con esa condición.

Ejemplo 1. Consideremos el grupo multiplicativo $\mathbf{C}^*$. Una raíz compleja p-ésima de la unidad, donde $p \geq 1$ es un entero, es por definición un $z \in \mathbf{C}^*$ tal que $z^p = 1$. En otros términos, el conjunto de las raíces complejas de la unidad, de orden no especificado, es el conjunto de los elementos periódicos de $\mathbf{C}^*$. Como es sabido, en álgebra elemental se define una *raíz primitiva* p-ésima de la unidad como cualquier $z \in \mathbf{C}^*$ que satisface la ecuación $z^p = 1$ y, en el caso $p \geq 2$, no satisface ninguna de las ecuaciones $z^n = 1$ ($n = 1, \ldots, p-1$). En otros términos, una raíz primitiva p-ésima de la unidad es un elemento de período igual a p en el grupo multiplicativo $\mathbf{C}^*$.

Ejemplo 2. Dado un plano euclideano P y un punto $O \in P$, la simetría $s: P \to P$ que a cada $x \in P$ le asocia su simétrico $s(x)$ (con relación a O), goza de la propiedad de que $s\{s(x)\} = x$ para todo $x \in P$, o sea $ss = I$, donde I es la transformación idéntica. De una manera más general, sea $f: E \to E$ una permutación de un conjunto E, tal que $f^2 = I$ (esto es, $ff = I$). La condición $f^2 = I$ es equivalente a $f^{-1} = f$, de modo que f es simplemente una permutación involutoria de E. Ahora, $f^2 = I$ nos dice que f es un elemento periódico del grupo $E!$ de las permutaciones de E y que su período p es un divisor de 2, esto es $p = 1$ o $p = 2$. El caso $p = 1$ se verifica sólo cuando $f = I$. En otros términos, las permutaciones involutorias de un conjunto E, distintas de la transformación idéntica, son los elementos de período 2 del grupo $E!$ En matemática elemental existen varios ejemplos, como las simetrías, de transformaciones de período 2.

En el caso de un grupo aditivo G, la noción de orden de un elemento $x \in G$ es introducida de manera similar a partir de la ecuación $nx = 0$ (en sustitución de $x^n = e$). El *orden* p es el entero natural caracterizado por las dos propiedades siguientes:

1. $px = 0$ y, de un modo más general, $mx = 0$ cualquiera que sea el múltiplo $m \in \mathbf{Z}$ de p.

2. si $mx = 0$, $m \in \mathbf{Z}$, entonces m es un múltiplo de p.

Los demás comentarios que se hicieron en el caso multiplicativo pueden ser repetidos aquí con las modificaciones pertinentes.

Ejemplo 3. Consideremos el grupo aditivo $\mathbf{Z}/p$ y, para fijar ideas, tomemos $p = 6$. En este grupo, el elemento I es periódico de período 6, pues las sumas $1 \overset{6}{+} 1$, $1 \overset{6}{+} 1 \overset{6}{+} 1, \ldots$ son todas diferentes de cero siempre que el número de sumandos sea menor que 6. El elemento 2 es de período 3, puesto que $2 \overset{6}{+} 2 = 4 \neq 0$, $2 \overset{6}{+} 2 \overset{6}{+} 2 = 0$. De igual forma, 3 tiene período 2, 4 tiene período 3 y 5 tiene período 6.

Las nociones de múltiplos y potencias dan lugar a una categoría de grupos extremadamente simple, a saber: los grupos cíclicos.

Dados un grupo aditivo G y $x \in G$, el conjunto H de los múltiplos enteros nx, $n \in \mathbf{Z}$, de x constituye un subgrupo de G. En efecto, $0 = 0x \in \in H$ y si mx y nx son dos elementos de H, entonces $mx - nx = (m - n)x$ será también un elemento de H. El grupo H tiene la propiedad de consistir de los múltiplos de un cierto elemento x. Un grupo aditivo se dice *cíclico* cuando contiene al menos un elemento x, cuyos varios múltiplos nx, $n \in \mathbf{Z}$ constituyen todo el grupo. Este x recibe el nombre de *generador* del grupo.

Las consideraciones precedentes se repiten trivialmente para los grupos multiplicativos. Se dice que un grupo multiplicativo es *cíclico* si contiene, por lo menos, un *generador*, cuyas potencias forman todo el grupo. Obsérvese que un grupo multiplicativo cíclico es, por fuerza, conmutativo. En efecto, si $y = x^m$ y $z = x^n$ son dos elementos del grupo (donde x es el generador considerado), entonces $yz = x^m x^n = = x^{m+n} = x^{n+m} = x^n x^m = zy$, como queríamos.

Vamos a mostrar ahora que los grupos cíclicos pueden clasificarse completamente salvo por isomorfismos. Para fijar la notación, consideremos el caso de un grupo aditivo G con generador x. Llamemos p al orden de x (pág. 26). Hay que distinguir los casos siguientes:

1. $p = 0$. Comencemos por recordar que $m, n \in \mathbf{Z}$, $m \neq n$, implican $mx \neq nx$, esto es, que los múltiplos de x son distintos dos a dos. Luego, G es infinito (o sea, contiene una infinidad de elementos). Afirmamos, entonces, que G es isomorfo al grupo aditivo $\mathbf{Z}$. En efecto, consideremos la correspondencia f de $\mathbf{Z}$ en G dada por $x \mapsto nx$. Observemos que f es un homomorfismo, pues $f(m + n) = (m + n)x = f(m) + f(n)$. Además, f es inyectiva, pues $m \neq n$ implica $mx \neq nx$, como ya notamos (o sea, $f(m) \neq f(n)$). Por último, f actúa sobre G, pues dado $y \in G$, la hipótesis de ser G cíclico permite escribir $y = nx$, esto es, $y = f(n)$. Esto completa la demostración de que f es un isomorfismo entre $\mathbf{Z}$ y G.

2. $p > 0$. Para comenzar, advirtamos que G consiste de los elementos (distintos dos a dos) siguientes: $0, x, 2x, \ldots, (p - 1)x$. En efecto, todo $y \in G$ es (por la hipótesis de ser G cíclico) de la forma $y = nx$, $n \in \mathbf{Z}$. Dividiendo n entre p y llamando q al cociente y r al residuo, tendremos $n = qp + r$, de donde $y = nx = (qp + r)x = (qp)x + rx = q(px) + rx = q0 + rx = rx$ y basta observar que $0 \leq r \leq p - 1$ para concluir que y es uno de los elementos $0, x, 2x, \ldots, (p - 1)x$. Además, estos elementos son distintos dos a dos, pues si tuviésemos $0 \leq m < n \leq p - 1$ y $mx = nx$, también tendríamos $(n - m)x = 0$. Por la definición de orden p de x, la diferencia $(n - m)$ debiera ser un múltiplo de p, lo que es imposible, ya que $0 < n - m \leq p - 1$. De ahí resulta, en particular, que G es finito y formado por p elementos. Afirmamos, entonces, que G es isomorfo al grupo aditivo $\mathbf{Z}/p$. De hecho, consideremos la correspondencia f de $\mathbf{Z}/p$ en G dada por $n \mapsto nx$. Esta f es un homomorfismo. Para mostrarlo, tomemos $m, n \in \mathbf{Z}/p$, dividamos $m + n$ entre p y sean q el cociente y r el residuo. Tenemos $m \overset{p}{+} n = r$, $m + n = qp + r$, de donde $(m + n)x = (qp + r)x = q(px) + rx = rx$, o sea $rx = mx + nx$, es decir $f(m \overset{p}{+} n) = f(m) + f(n)$, como queríamos. Además, f es inyectiva, pues $m, n \in \mathbf{Z}/p$, $m \neq n$, implica $mx \neq nx$ (pues los elementos $0, x, 2x, \ldots, (p - 1)x$ son distintos dos a dos), o $f(m) \neq f(n)$. Finalmente, f actúa sobre G, pues, como ya dijimos, todo $y \in G$ es de la forma $y = nx$, $n \in \mathbf{Z}/p$, o $y = f(n)$. Queda así completa la demostración de que f es un isomorfismo entre $\mathbf{Z}/p$ y G.

Como estos dos casos son los únicos a considerar (salvo por la cuestión de notación aditiva o multiplicativa), podemos resumir la discusión precedente del siguiente modo:

Proposición 3. Todo grupo cíclico infinito es isomorfo al grupo aditivo $\mathbf{Z}$. Todo grupo cíclico finito con $p \geq 1$ elementos es isomorfo al grupo aditivo $\mathbf{Z}/p$.

Ejemplo 4. Consideremos el grupo multiplicativo de las raíces complejas p-ésimas de la unidad (pág. 64). Se sabe, del álgebra elemental, que sólo existen p raíces p-ésimas de la unidad, a saber:

$$e^{2\pi n i/p} \quad (n = 0,\ 1,\ 2, \ldots, p - 1).$$

Si se escribe $x = e^{2\pi i/p}$, vemos que los varios elementos de este grupo son las potencias $x^n (n = 0, 1, 2, \ldots, p - 1)$ de x, lo que prueba que el grupo es cíclico con p elementos y que x es uno de sus generadores.

Ejercicios

1. En todo grupo conmutativo, el conjunto de los elementos periódicos es un subgrupo. Mostrar, con un ejemplo, que lo mismo puede ser verdad aun en la ausencia de conmutatividad.

2. Todo elemento de un grupo finito es periódico.

3. Un grupo cíclico infinito sólo admite dos generadores. Dos grupos cíclicos infinitos siempre son isomorfos y entre ellos existen sólo dos isomorfismos posibles.

4. Un grupo cíclico finito con $p \geq 1$ elementos, admite $\varphi(p)$ generadores, donde $\varphi(p)$ indica el número de términos de la secuencia $1, 2, \ldots, p - 1$ que son primos con p si $p \geq 2$ y $\varphi(1) = 1$ (φ recibe el nombre de *función de Euler*). Dos grupos cíclicos finitos con p elementos son siempre isomorfos y entre ellos existen $\varphi(p)$ isomorfismos posibles.

5. El orden de todo elemento de un grupo cíclico finito G divide al número de elementos del grupo.

3

ANILLOS

La teoría de los anillos nació del estudio de temas relacionados con la divisibilidad entre enteros, del estudio paralelo de la divisibilidad entre polinomios y de la teoría de cuerpos tales como la de los números racionales, reales, complejos, algebraicos, de los cuaterniones, de las fracciones racionales, de las funciones algebraicas, etc. Al principio, fueron los problemas de la teoría de números y de la geometría algebraica los que dieron lugar a los conceptos de anillo, cuerpo e ideal. En su forma axiomática, tales nociones fueron concebidas por Dedekind y otros a fines del siglo pasado. Sus aplicaciones al análisis, que reflejan las tendencias recientes de algebrización de esta rama de las matemáticas, datan apenas del segundo cuarto de nuestro siglo.

§ 1. ANILLOS CONMUTATIVOS

En muchos de los conjuntos sobre los que trata el álgebra elemental es posible sumar y multiplicar dos elementos del conjunto. A título de ejemplos sencillos, consideremos los casos siguientes:

Ejemplo 1. En el conjunto **Z** de los enteros tenemos dos operaciones, una de adición y otra de multiplicación, que a cada par ordenado (x, y) asocian, respectivamente, un elemento $x + y \in \mathbf{Z}$ llamado la suma de x e y, y otro $xy \in \mathbf{Z}$, que recibe el nombre de producto. Con respecto a la suma, **Z** es un grupo aditivo (págs. 47-48). Asimismo, la multiplicación es conmutativa, asociativa y distributiva con respecto a la suma, esto es

$$xy = yx, \quad x(yz) = (xy)z, \quad x(y + z) = xy + xz.$$

Notemos que **Z** no constituye un grupo multiplicativo con respecto a la multiplicación (¿por qué?). Análogamente, cada uno de los conjuntos **Q**, **R** y **C** (pág. 1) posee las propiedades que acaban de ser indicadas para **Z** en relación con las operaciones de suma y multiplicación habituales. Resaltemos apenas una diferencia entre el caso de **Z**, por una parte, y el de **Q**, **R** y **C** por la otra: el conjunto $\mathbf{Q}^*$ de los números racionales diferentes de 0 constituye un grupo multiplicativo con respecto a la multiplicación usual (pág. 51), y lo mismo puede decirse de los conjuntos $\mathbf{R}^*$ y $\mathbf{C}^*$ (pág. 52). Sin embargo, el conjunto $\mathbf{Z}^*$ de los enteros no nulos, no es un grupo multiplicativo (¿por qué?). Esta observación será más tarde reforzada con la distinción entre anillos y cuerpos.

Ejemplo 2. Consideremos el conjunto F de las funciones reales de variable real definidas en un intervalo $[a, b]$. Ya hemos definido la suma $f + g \in F$ de dos funciones f, $g \in F$ (pág. 46). Añadamos ahora la definición del producto $fg \in F$ como la función tal que

$$(fg)(x) = f(x)g(x)$$

para $x \in [a, b]$, o sea fg es la función que en el punto x toma el valor $f(x)g(x)$. Tenemos ahora dos operaciones en F, una de suma y otra de multiplicación. Sabemos ya (pág. 47) que F es un grupo aditivo con respecto a esta suma. Además, afirmamos que la multiplicación en F es conmutativa, asociativa y distributiva con respecto a la suma. En efecto,

$$(fg)(x) = f(x)g(x), \quad (gf)(x) = g(x)f(x),$$

de donde $(fg)(x) = (gf)(x)$, para todo $x \in [a, b]$, o sea $fg = gf$. Además,

$$[f(gh)](x) = f(x)[(gh)(x)] = f(x)[g(x)h(x)],$$

$$[(fg)h](x) = [(fg)(x)]h(x) = [f(x)g(x)]h(x),$$

de donde $[f(gh)](x) = [(fg)h](x)$ para todo $x \in [a, b]$, lo que significa $f(gh) = (fg)h$. Por último,

$$[f(g+h)](x) = f(x)[(g+h)(x)] = f(x)[g(x) + h(x)],$$

$$(fg + fh)(x) = (fg)(x) + (fh)(x) = f(x)g(x) + f(x)h(x),$$

o sea $[f(g+h)](x) = (fg + fh)(x)$, para todo $x \in [a, b]$, esto es $f(g+h) = fg + fh$.

La analogía entre el conjunto F dotado de las operaciones de suma y multiplicación y los conjuntos **Z**, **Q**, **R** y **C** es patente.

Notemos que el hecho de que las funciones reales que constituyen F estén definidas sólo en un intervalo cerrado $[a, b]$, o definidas en la recta **R** o hasta en un cierto conjunto E arbitrario, no altera en nada la esencia de este ejemplo.

Ejemplo 3. Designemos con $\mathbf{R}[x]$ al conjunto de los polinomios reales de una variable real x (pág. 2). Si p y q, dados por

$$p(x) = a_0 + a_1x + \ldots + a_mx^m, \quad q(x) = b_0 + b_1x + \ldots + b_nx^n$$

fueran dos polinomios, definamos la suma de p y q por

$$(p+q)(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \ldots,$$

o sea $p + q$ es el polinomio que se obtiene sumando los términos del mismo grado en $p(x)$ y $q(x)$. Definamos también el producto pq por

$$(pq)(x) = a_0b_0 + (a_0b_1 + a_1b_0)x + (a_0b_2 + a_1b_1 + a_2b_0)x^2 + \ldots + a_mb_nx^{m+n}$$

o sea pq es el polinomio obtenido al multiplicar cada término de $p(x)$ por cada término de $q(x)$ y reducir términos semejantes. Como es sabido por álgebra elemental, el conjunto $\mathbf{R}[x]$ posee, con respecto a la adición y a la multiplicación, todas las propiedades de los ejemplos

precedentes. Por ejemplo, además de la expresión para $(pq)(x)$ presentada antes, tenemos también que $(qp)(x) = b_0a_0 + (b_0a_1 + b_1a_0)x + (b_0a_2 + b_1a_1 + b_2a_0)x^2 + \ldots + b_na_mx^{n+m}$, de donde $(pq)(x) = (qp)(x)$ para cualquier x real o sea que $pq = qp$. Del mismo modo se prueban todas las otras propiedades. Sin embargo, el modo más cómodo de verificarlas consiste en reducirlas al ejemplo anterior. En efecto, notemos que por ser $(p + q)(x) = (a_0 + b_0) + (a_1 + b_1)x + \ldots$, se tiene que $(p + q)(x) = p(x) + q(x)$; en otros términos, la suma $p + q$ de dos polinomios es precisamente la que se obtiene pensando en p y q como funciones de la variable real x y procediendo como en el ejemplo 2. Lo mismo para el producto, pues de $(pq)(x) = a_0b_0 + (a_1b_0 + a_0b_1)x + \ldots$ resulta que $(pq)(x) = p(x)q(x)$. Ahora bien, como las leyes conmutativa, asociativa, etc. son satisfechas entre funciones cualesquiera (ejemplo 2), ellas son en particular satisfechas entre polinomios. Un pequeño comentario: uno de los motivos por los cuales es más cómodo verificar las leyes conmutativa, asociativa, etc. para polinomios a partir de las mismas leyes para las funciones es que la expresión del último término en $(p + q)(x) = (a_0 + b_0) + (a_1 + b_1)x + \ldots$ depende del hecho de ser $m < n$, $m = n$, o bien $m > n$ (donde m y n son los grados de p y q).

Lo que se acaba de indicar para $\mathbf{R}[x]$ se repite sin mayores comentarios para el conjunto de los polinomios complejos de una variable compleja x, que se representará por $\mathbf{C}[x]$.

En aritmética elemental se estudian las propiedades de las operaciones de adición y multiplicación en $\mathbf{Z}$ y $\mathbf{Q}$, en tanto que el estudio paralelo de $\mathbf{R}$, $\mathbf{C}$, $\mathbf{R}[x]$ y $\mathbf{C}[x]$ es objeto del álgebra elemental. Por ejemplo, en aritmética se estudian las cuestiones de divisibilidad y máximo y mínimo común múltiplos entre enteros, en tanto que en álgebra se hace un estudio semejante para polinomios. Si se quiere sintetizar estos varios aspectos comunes en una sola teoría es necesario introducir los llamados anillos.

Un *anillo conmutativo* es un conjunto A, donde están dadas dos operaciones, una de adición y la otra de multiplicación, que satisfacen las condiciones siguientes:

1. En relación con la adición, A es un grupo aditivo (son, pues, satisfechas las condiciones 1, ..., 5 de las págs. 47-48).

2. La operación de multiplicación es una función definida en $A \times A$, con valores en A, que a cada par ordenado (x, y) de $A \times A$ le asocia un elemento $xy \in A$, llamado el *producto de los factores* x e y.

3. La multiplicación es conmutativa, esto es:

$$xy = yx \qquad (x, y \in A).$$

4. La multiplicación es asociativa:

$$x(yz) = (xy)z \qquad (x, y, z \in A).$$

5. La multiplicación es distributiva con respecto a la adición:

$$x(y + z) = xy + xz \quad (x, y, z \in A).$$

Los varios conjuntos **Z**, **Q**, **R**, **C**, F, $\mathbf{R}[x]$ y $\mathbf{C}[x]$, arriba mencionados, proveen ejemplos sencillos de anillos conmutativos. Por conveniencia, se acostumbra también representar el producto en un anillo por $x \cdot y$, $x \times y$, etc.

Cerraremos esta sección mencionando un ejemplo de anillo *finito*, esto es con un número finito de elementos.

Ejemplo 4. Dado el entero $p \geq 1$, se sabe ya que el conjunto $\mathbf{Z}/p$ constituido por los enteros 0, 1, ..., $p - 1$ constituye un grupo aditivo con relación a la adición módulo p (págs. 49-50). Vamos ahora a extender la definición de multiplicación módulo p y constatar que $\mathbf{Z}/p$ se vuelve, entonces, un anillo conmutativo. Si $x, y \in \mathbf{Z}/p$, definamos su *producto módulo* p como el residuo de la división entre p del producto usual xy. Usaremos el símbolo $x \overset{p}{\cdot} y$ para designar este nuevo producto. Es claro que la multiplicación módulo p es conmutativa, pues $x \overset{p}{\cdot} y$ e $y \overset{p}{\cdot} x$ son los restos de la división entre p de xy e yx, respectivamente, y $xy = yx$. La ley asociativa

$$x \overset{p}{\cdot} (y \overset{p}{\cdot} z) = (x \overset{p}{\cdot} y) \overset{p}{\cdot} z$$

se verifica exactamente como en el caso de la adición módulo p (pág. 50), pues los dos miembros precedentes son iguales al residuo del producto usual xyz dividido entre p. Verifiquemos la ley distributiva

82

$$x \overset{p}{\cdot} (y \overset{p}{+} z) = x \overset{p}{\cdot} y \overset{p}{+} x \overset{p}{\cdot} z$$

Para obtener $y \overset{p}{+} z$, debemos dividir $(y + z)$ entre p. Sean q el cociente y r el residuo, tenemos

$$y \overset{p}{+} z = r, \qquad y + z = pq + r.$$

Además, para calcular $x \overset{p}{\cdot} (y \overset{p}{+} z) = x \overset{p}{\cdot} r$, debemos dividir xr entre p. Llamando q' al cociente y r' al residuo, tenemos

$$x \overset{p}{\cdot} (y \overset{p}{+} z) = r', \qquad xr = pq' + r'.$$

En consecuencia,

$$x(y + z) = xpq + xr = xpq + pq' + r' = p(xq + q') + r',$$

igualdad que prueba que r' es el residuo de la división de $x(y + z)$ entre p:

$$x \overset{p}{\cdot} (y \overset{p}{+} z) = \text{residuo de dividir } x(y + z) \text{ entre } p.$$

Por otro lado, para calcular $x \overset{p}{\cdot} y$ y $x \overset{p}{\cdot} z$, hay que dividir xy y xz entre p. Sean Q y Q' los cocientes y R y R' los residuos, se tiene

$$x \overset{p}{\cdot} y = R, \; xy = pQ + R; \; x \overset{p}{\cdot} z = R', \; xz = pQ' + R'.$$

Además, para calcular $x \overset{p}{\cdot} y \overset{p}{+} x \overset{p}{\cdot} z = R \overset{p}{+} R'$, debemos dividir $R + R'$ entre p, lo que da un cociente Q'' y un residuo R'':

$$x \overset{p}{\cdot} y \overset{p}{+} x \overset{p}{\cdot} z = R'', \qquad R + R' = pQ'' + R''.$$

De ahí resulta que

$$xy + xz = (pQ + R) + (pQ' + R') = p(Q + Q') + (R + R') =$$

$$= p(Q + Q' + Q'') + R'',$$

igualdad que prueba que R'' es el residuo de la división entre p de $xy + xz$:

$$x \overset{p}{\cdot} y \overset{p}{+} x \overset{p}{\cdot} z = \text{residuo entre } p \text{ de } xy + xz.$$

Como $x(y + z) = xy + xz$, se concluye que la multiplicación módulo p es distributiva con relación a la adición módulo p, como se quería probar.

Ejercicios

1) Sean A un grupo aditivo y 0 su elemento cero, definamos una multiplicación en A mediante $xy = 0$ para cualesquiera $x, y \in A$. Constatar que A es, entonces, un anillo conmutativo. (Un anillo tal, en el que el producto de dos elementos cualesquiera es siempre cero, es llamado un anillo *trivial*).

§ 2. ANILLOS ARBITRARIOS

Todos los anillos que habitualmente son mencionados como tales en los cursos elementales de álgebra son conmutativos. Se encuentran, empero, ejemplos importantes de conjuntos cuyos elementos pueden sumarse y multiplicarse; la adición es conmutativa sin que la multiplicación lo sea. De ahí la necesidad del concepto más general de anillo no necesariamente conmutativo.

Un *anillo* es un conjunto A, donde están dadas dos operaciones, una de adición y otra de multiplicación, tales que:

1. Con respecto a la adición, A es un grupo aditivo.

2. La multiplicación, definida de $A \times A$ en A, asocia a cada par ordenado (x, y) un producto $xy \in A$.

3. La multiplicación es asociativa:

$$x(yz) = (xy)z \qquad (x, y, z \in A).$$

4. La multiplicación es doblemente distributiva con respecto a la suma:

$$x(y + z) = xy + xz, \quad (y + z)x = yx + zx \quad (x, y, z \in A).$$

En pro de la claridad, resaltemos que la condición 1 engloba, entre otras, a la condición $x + y = y + x$, pero $xy = yx$ no se satisface necesariamente, como muestra el siguiente ejemplo.

Ejemplo 1. Sean x_1 y x_2 dos variables reales. Una función lineal homogénea, o forma lineal, en x_1 y x_2 es toda expresión del tipo $a_1x_1 + a_2x_2$, donde a_1 y a_2 son dos coeficientes reales dados. Supongamos que dos variables reales y_1 e y_2 estén ligadas a x_1 y x_2 por las ecuaciones

$$y_1 = a_{11}x_1 + a_{12}x_2,$$

$$y_2 = a_{21}x_1 + a_{22}x_2,$$

esto es, y_1 e y_2 son formas lineales en x_1 y x_2. Como se sabe, se da el nombre de *matriz* del sistema anterior al conjunto de los cuatro coeficientes a_{ij} escritos ordenadamente tal como se presentan. Indicando esta matriz con a, tenemos

$$a = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Representaremos con $\mathbf{M}_2(\mathbf{R})$ al conjunto de las matrices reales de orden 2, esto es con dos líneas y dos columnas y elementos a_{ij} reales. La noción de matriz proviene del estudio de los sistemas lineales. Sin embargo, es importante aprender a pensar en matrices independientemente de los sistemas lineales que las determinan: éstos se mencionan aquí sólo como motivación del concepto de matriz y de las definiciones de suma y producto de matrices. Si, además de las relaciones arriba citadas entre y_1, y_2 y x_1, x_2, tuviéramos las relaciones siguientes entre ciertas variables z_1, z_2 y x_1, x_2,

$$z_1 = b_{11}x_1 + b_{12}x_2,$$

$$z_2 = b_{21}x_1 + b_{22}x_2,$$

con matriz

$$b = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

entonces, escribiendo $t_1 = y_1 + z_1$, $t_2 = y_2 + z_2$, las nuevas variables t_1 y t_2 se expresan por medio de x_1 y x_2 del siguiente modo:

$$t_1 = (a_{11} + b_{11})x_1 + (a_{12} + b_{12})x_2$$

$$t_2 = (a_{21} + b_{21})x_1 + (a_{22} + b_{22})x_2.$$

Por tanto, es natural definir la suma $a + b$ de las matrices a y b por

$$a + b = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{pmatrix}$$

esto es $a + b$ es la matriz obtenida sumando los elementos correspondientes de a y b. Supongamos, ahora, que además de las relaciones

arriba escritas entre x_1, x_2 e y_1, y_2, tenemos dos nuevas variables t_1 y t_2 vinculadas a x_1 y x_2 por las relaciones

$$x_1 = b_{11}t_1 + b_{12}t_2,$$

$$x_2 = b_{21}t_1 + b_{22}t_2$$

de matriz b. Sustituyendo, entonces, las expresiones de x_1, x_2 en función de t_1 y t_2 en las ecuaciones que dan y_1 e y_2, se tendrá

$$y_1 = (a_{11}b_{11} + a_{12}b_{21})t_1 + (a_{11}b_{12} + a_{12}b_{22})t_2,$$

$$y_2 = (a_{21}b_{11} + a_{22}b_{21})t_1 + (a_{21}b_{12} + a_{22}b_{22})t_2,$$

por este motivo, definimos el producto ab de las dos matrices a y b por

$$ab = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

esto es ab es la matriz cuyo elemento en la fila i y columna j se obtiene multiplicando ordenadamente los elementos de la fila i en a por los elementos de la columna j en b y sumando los resultados. Lo que se acaba de indicar para las matrices de orden 2 se repite para las matrices de orden n, esto es de n filas y n columnas. Se acostumbra representar por (a_{ij}) a la matriz cuyo elemento genérico es a_{ij}, sobrentendiéndose que el primer índice (i) indica la fila, y el segundo (j) la columna de ese elemento, i y j deben variar de 1 a n. Si $a = (a_{ij})$ y $b = (b_{ij})$, se definen la suma $a + b$ y el producto ab por

$$a + b = (a_{ij} + b_{ij}), \qquad ab = \left(\sum_{k=1}^{n} a_{ik}b_{kj}\right).$$

Afirmamos, pues, que con relación a la adición y multiplicación que acaban de definirse, el conjunto $\mathrm{M}_n(\mathbf{R})$ de las matrices cuadradas de orden n con elementos reales es un anillo. A título de ejemplo, verifiquemos la ley asociativa $(ab)c = a(bc)$, donde $a = (a_{ij})$, $b = (b_{ij})$ y $c = (c_{ij})$ son tres elementos arbitrarios de $\mathrm{M}_n(\mathbf{R})$. Aplicando la definición de producto, se tiene $ab = (r_{ij})$, donde

$$r_{ij} = \sum_k a_{ik}b_{kj}.$$

Por tanto, $(ab)c = (s_{ij})$, donde

$$s_{ij} = \sum_h r_{ih}c_{hj} = \sum_h \left(\sum_k a_{ik}b_{kh}\right)c_{hj} = \sum_{h,k} a_{ik}b_{kh}c_{hj}.$$

De un modo perfectamente análogo se halla que el elemento en la fila i y columna j de $a(bc)$ es también

$$\sum_{h,k} a_{ik}b_{kh}c_{hj},$$

lo que prueba la igualdad $(ab)c = a(bc)$. El resto de la demostración de que $\mathbf{M}_n(\mathbf{R})$ es un anillo es cuestión de rutina y se deja a cargo del lector. Adviértase solamente que en el anillo $\mathbf{M}_n(\mathbf{R})$ el elemento cero es la matriz de orden n, cuyos elementos son todos iguales a cero. Por ejemplo, en el caso $n = 2$, tenemos

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} a_{11} + 0 & a_{12} + 0 \\ a_{21} + 0 & a_{22} + 0 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

lo que justifica nuestra aserción. Además, la simétrica $-a$ de una matriz a es la matriz obtenida tomando los simétricos $-a_{ij}$ de los varios elementos a_{ij} de a, cuya demostración es inmediata.

Por último, mostremos que $\mathbf{M}_n(\mathbf{R})$, $n \geq 2$, es un anillo no conmutativo. Nos limitaremos a $n = 2$. Sean

$$a = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad b = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

dos matrices, la segunda de las cuales vamos a determinar. Se ve de inmediato que

$$ab = \begin{pmatrix} b_{11} & b_{12} \\ 0 & 0 \end{pmatrix} \qquad ba = \begin{pmatrix} b_{11} & 0 \\ b_{21} & 0 \end{pmatrix}$$

Es, pues, extremadamente fácil escoger los elementos de la matriz b de modo que $ab \neq ba$: por ejemplo, basta tomar $b_{11} = 0$, $b_{12} = 1$, $b_{21} = 0$, $b_{22} = 0$, o sea

$$b = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Lo que acabamos de hacer para las matrices reales se repite sin dificultad para el conjunto $\mathbf{M}_n(\mathbf{C})$ de las matrices complejas (a_{ij}) de orden n, esto es con elementos $a_{ij} \in \mathbf{C}$.

En todo anillo, las leyes asociativas de la adición y de la multiplicación permiten definir sumas y productos de tres o más elementos

$$x + y + z = x + (y + z) = (x + y) + z, \qquad xyz = x(yz) = (xy)z,$$

exactamente como en los casos de los grupos aditivos y multiplicativos. Se emplearán los signos Σ y Π para abreviar sumas y productos. Finalmente, se aplicarán las mismas reglas de omisión o inserción de paréntesis del álgebra elemental. Por ejemplo, $xy + z$ significa $(xy) + z$. Otro ejemplo: $-xy$ representa $-(xy)$, etc.

Diremos que dos elementos x e y de un anillo *conmutan* si $xy = yx$. Un *anillo conmutativo* es, precisamente, todo anillo donde dos elementos cualesquiera conmutan.

Proposición 1. En todo anillo A se tiene

1. $0x = x0 = 0$, 2. $-xy = (-x)y = x(-y)$, 3. $(-x)(-y) = xy$,

4. $x(y - z) = xy - xz$, $(y - z)x = yx - zx$.

Demostración. Tenemos que $0x = (0 + 0)x = 0x + 0x$, de donde $0x = 0$ (por el ítem 3, prop. 1, pág. 49). Análogamente para $x0 = 0$.

Además, $xy + (-x)y = [x + (-x)]\, y = 0y = 0$, lo que prueba que $-xy = (-x)y$. La demostración de que $-xy = x(-y)$ es similar.

Notemos ahora que $(-x)(-y) = -[x(-y)] = -(-xy) = xy$ (por el ítem 1, prop. 1, pág. 49).

Finalmente, $x(y - z) = x[y + (-z)] = xy + x(-z) = xy + (-xz) = xy - xz$. De modo análogo, para $(y - z)\,x = yx - zx$. QED

Obsérvese que en todo anillo, la ley distributiva es válida también para una suma algebraica de tres o más sumandos. Por ejemplo,

$$t(x - y + z) = tx - ty + tz,$$

Como es fácil verificar.

Un anillo A es, en particular, un grupo aditivo, Por tanto, para todo entero $n \in \mathbf{Z}$ y todo $x \in A$ tiene sentido el múltiplo $nx \in A$ (pág. 73). No se debe confundir tal producto de un entero por un elemento del anillo con el producto de dos elementos del anillo: en este último, los dos factores pertenecen al anillo, en tanto que en el primero, uno de los factores es un entero y el otro es un elemento del anillo. 87

Proposición 2. Si $n \in \mathbf{Z}$ y x, $y \in A$, entonces

$$n(xy) = (nx)y = x(ny).$$

Demostración. Si $n \geq 1$, se tiene

$$n(xy) = xy + \ldots + xy\,(n \text{ veces}) = (x + \ldots + x)\,y\,(n \text{ veces}) = (nx)y.$$

Si $n = 0$, notemos que $0(xy) = 0$ y que $(0x)y = 0y = 0$. Finalmente, si $n < 0$, caso en el que $n = -N$ con $N \geq 1$, se tiene

$$n(xy) = (-N)\,(xy) = -N(xy) = -(Nx)y = (-Nx)y = (nx)y.$$

La demostración de que $n(xy) = x(ny)$ es análoga. QED

Si $n = 1, 2, 3, \ldots$, y $x \in A$, se puede definir la potencia x^n por medio de $x^1 = x$, $x^2 = xx$, $x^3 = xxx$, ... Al contrario de lo que vimos para los grupos multiplicativos (pág. 74), no se define en el caso general la potencia x^n para $n \leq 0$ entero. Más tarde se volverá a tratar este caso.

Proposición 3. Si m, $n = 1, 2, \ldots$ y x, $y \in A$, se tiene que

1. $x^{m+n} = x^m x^n$, 2. $(xy)^n = x^n y^n$ si es que $xy = yx$,

3. $(x^n)^m = x^{mn}$, 4. $x^1 = x$.

La demostración es inmediata y será, por tanto, omitida.

Ejercicios

1) Sea $\mathcal{F}$ el conjunto de las funciones reales de variable real. Definamos en $\mathcal{F}$ una operación de adición y otra de multiplicación mediante $(f + g)(x) = f(x) + g(x)$ y $(fg)(x) = f\{g(x)\}$, donde f, $g \in \mathcal{F}$ y $x \in \mathbf{R}$. Mostrar que $\mathcal{F}$ satisface todas las condiciones de la noción de anillo, menos una.

2) En cálculo vectorial del espacio euclideano tridimensional se considera el conjunto $\mathcal{V}$ de los segmentos orientados, o vectores, de origen dado O y se definen la operación de adición (por la regla del paralelogramo) y la multiplicación vectorial. Mostrar que $\mathcal{V}$ satisface todas las condiciones del concepto de anillo, menos una.

3) En un anillo se tiene $(-x)^n = x^n$ o $(-x)^n = -x^n$ según que el entero $n \geq 1$ sea par o impar. Mostrar que la fórmula del binomio de Newton

$$(x + y)^n = \sum_{i=1}^{n} \binom{n}{i} x^i y^{n-i}$$

es válida siempre que $xy = yx$.

4) Sea A un anillo con suma $x + y$ y producto xy. Manteniendo la definición de suma, alteremos la definición de producto, escribiendo $x \cdot y = -xy$. Demostrar que, con respecto a las operaciones $(x, y) \mapsto x + y$, y $(x, y) \to x \cdot y$, A es un anillo.

5) Consideremos el anillo $\mathbf{Z}$ con respecto a la suma $x + y$ y al producto xy usuales. Para que $\mathbf{Z}$ sea un anillo con respecto a la suma $x + y$ y a un nuevo producto $x \cdot y$, es necesario y suficiente que exista un $a \in \mathbf{Z}$ tal que $x \cdot y = axy$, donde el producto axy se entiende en el sentido usual. Un hecho similar vale para el anillo $\mathbf{Q}$. Idem para $\mathbf{Z}/p$, donde ahora $x \cdot y = a \overset{p}{\cdot} x \overset{p}{\cdot} y$.

§ 3. SUBANILLOS

La noción de subanillo es a la teoría de los anillos lo que la noción de subgrupo es a la teoría de grupos, o la de subconjunto es a la teoría de conjuntos. Con pocas excepciones, cuando se consideran dos anillos A y B, y B está contenido en A, las operaciones de adición y multiplicación en B y A se comportan del mismo modo; más explícitamente, si x, $y \in B$ (de donde resulta que x, $y \in A$), la suma $x + y$ y el producto xy tienen los mismos valores cuando se calculan considerando a x e y como elementos de B o de A. Es lo que sucede, por ejemplo, en el caso de $\mathbf{Z}$ y $\mathbf{Q}$, donde $\mathbf{Z} \subset \mathbf{Q}$. Uno de los pocos ejemplos habituales en que

las operaciones de B y A actúan de modo diferente en los dos anillos es el de $\mathbf{Z}/p$ y $\mathbf{Z}$, donde $\mathbf{Z}/p \subset \mathbf{Z}$.

Un subconjunto B de un anillo A se dice un *subanillo* de A cuando B es un anillo con respecto a las operaciones de adición y multiplicación de A restringidas a los elementos de B. Esto quiere decir que B es un anillo con respecto a las dos correspondencias $(x, y) \mapsto x + y$ y $(x, y) \mapsto xy$, donde $x, y \in B$ y $x + y$ y xy se calculan como en A.

Proposición 1. Para que el subconjunto B del anillo A sea un subanillo es necesario y suficiente que:

1. B sea un subgrupo aditivo de A.

2. $x, y \in B$ implique $xy \in B$.

Demostración. Supongamos satisfechas las condiciones 1 y 2 mencionadas. Por la primera condición, vemos que B es un grupo aditivo con respecto a la adición de A restringida a los elementos de B. Por la condición 2, vemos que la correspondencia $(x, y) \mapsto xy$, donde $x, y \in B$ y xy se calcula en el supuesto de que x e y son elementos de A, actúa de $B \times B$ en B. Ahora bien, como la ley asociativa se satisface entre los elementos de A, en particular se satisface entre los elementos de B. Lo mismo es válido para las dos leyes distributivas. Luego, B es un anillo. Notemos que si A fuera conmutativo, lo mismo sucedería con B.

Recíprocamente, supongamos que B sea un subanillo de A. En particular, B es un subgrupo aditivo de A, esto es, se satisface 1. Además, la correspondencia $(x, y) \mapsto xy$, donde $x, y \in B$ y xy se calcula en el supuesto de que x e y son elementos de A, debe actuar de $B \times B$ en B. Esto exige que $xy \in B$, lo que prueba la mencionada condición 2. QED

Recordemos que en las aplicaciones de la proposición anterior se demuestra la condición 1 gracias a las proposiciones 1 ó 2 de las páginas 57 y 59.

Ejemplo 1. $\mathbf{Z}$ es un subanillo de $\mathbf{Q}$. De modo análogo, $\mathbf{Q}$ es un subanillo de $\mathbf{R}$ y $\mathbf{R}$ es un subanillo de $\mathbf{C}$.

El conjunto de los números complejos algebraicos (págs. 58-59) es un subanillo de $\mathbf{C}$. En efecto, ya sabemos que este conjunto es un grupo aditivo de $\mathbf{C}$. Sean x_1 e y_1 dos números algebraicos que satisfacen, respectivamente, las ecuaciones.

$$x^m + a_1 x^{m-1} + \ldots + a_m = 0, \qquad y^n + b_1 y^{n-1} + \ldots + b_n = 0,$$

cuyos coeficientes son números racionales. Además de x_1, la primera ecuación posee raíces $x_2, \ldots, x_m$. De la misma manera, además de y_1, la segunda ecuación posee raíces $y_2, \ldots, y_n$. Como se demuestra en la teoría elemental de las ecuaciones algebraicas, los números $x_i y_j$ $(i = 1, \ldots, m;\ j = 1, \ldots, n)$ son las raíces de una ecuación

$$z^{mn} + c_1 z^{mn-1} + \ldots + c_{mn} = 0$$

cuyos coeficientes $c_1, \ldots, c_{mn}$ se expresan como polinomios de $a_1, \ldots, a_m$ y $b_1, \ldots, b_n$ y, por consiguiente, también son números racionales. Como $x_i y_i$ es raíz de esta ecuación de coeficientes racionales, vemos que el producto de dos números algebraicos es también algebraico, como era necesario verificar. Luego, los números complejos algebraicos constituyen otro ejemplo de anillo.

Ejemplo 2. Consideremos el anillo F de las funciones reales definidas en $[a, b]$ (pág. 79). Afirmamos que el conjunto C de las funciones reales continuas en $[a, b]$ es un subanillo de F. En efecto, ya sabemos (pág. 59) que este conjunto es un subgrupo aditivo de F. Además, es sabido que el producto fg de dos funciones f, $g \in F$ continuas es también una función continua. Luego, las funciones reales continuas en $[a, b]$ constituyen un anillo, a saber: un subanillo de F.

Otro ejemplo semejante de anillo de funciones está dado por las funciones reales derivables en $[a, b]$.

Ejemplo 3. En el anillo $\mathbf{R}[x]$ (págs. 80-81), el subconjunto formado por los polinomios pares, esto es de la forma

$$p(x) = a_0 + a_2x^2 + a_4x^4 + \ldots + a_{2m}x^{2m},$$

constituye a las claras un subanillo. Lo mismo no sucede con el conjunto de los polinomios *impares*, esto es de la forma

$$q(x) = b_1x + b_3x^3 + b_5x^5 + \ldots + b_{2n-1}x^{2n-1}.$$

En el caso de un grupo aditivo, ya definimos las notaciones $X + Y$, $-X$ y $X - Y$, donde X y Y, son las dos partes cualesquiera del grupo (pág. 59). Si A es un anillo, definamos también

$$XY = \{xy;\ x \in X,\ y \in Y\},$$

o sea XY es el conjunto de los elementos de A de la forma xy, donde x e y varían en X e Y, respectivamente (exactamente como en el caso de los grupos multiplicativos, pág. 60). Es posible entonces reformular la proposición 1 del siguiente modo:

Proposición 2. Para que el subconjunto B del anillo A sea un subanillo es necesario y suficiente que:

1. $0 \in B$; 2. $B + B \subset B$; 3. $-B \subset B$; 4. $BB \subset B$.

Si se desea (pág. 60), se pueden sustituir en este enunciado las condiciones 2 y 3 por $B - B \subset B$.

Uno de los aspectos importantes del estudio de un anillo consiste en la descripción de sus subanillos. Al contrario de lo que sucede en la mayoría de los anillos habituales, en el caso de Z esta descripción es muy fácil, según se desprende de la proposición 5, página 64 y del resultado siguiente:

Proposición 3. Todo subgrupo aditivo de $\mathbf{Z}$ es también un subanillo de $\mathbf{Z}$ y recíprocamente.

Demostración. Sea H un subgrupo aditivo de $\mathbf{Z}$. Se sabe ya (pág. 64) que existe un entero, y sólo uno, $p \geq 0$ tal que H es el conjunto de los múltiplos enteros de p. Por tanto, si x e $y \in H$, donde $x = mp$ e $y = np$ (con m, $n \in \mathbf{Z}$), entonces $xy = nmp^2 = (mnp)p$, esto es xy es múltiplo de p, de donde $xy \in H$. Por la proposición 1 de esta sección, vemos que H es un subanillo de $\mathbf{Z}$. Recíprocamente, todo subanillo de $\mathbf{Z}$ (como todo subanillo de cualquier anillo) es también un grupo aditivo, conforme a la condición 1 de la proposición 1 de esta sección. QED

Así, vemos que en el caso del anillo $\mathbf{Z}$, no hay distinción entre subgrupos aditivos y subanillos de $\mathbf{Z}$. No se debe inferir por ello que lo mismo sea válido para otros anillos. Por ejemplo, en el caso del anillo $\mathbf{Q}$, el conjunto H de los números racionales de la forma $m/2$, donde m es entero, es un subgrupo de $\mathbf{Q}$ (pues $0 = 0/2$, $m/2 + n/2 = (m + n)/2$, $-(m/2) = (-m)/2$), el cual no es un subanillo (pues $1/2 \in H$, pero $1/2 \times 1/2 \notin H$).

Ejercicios

1) Todo subgrupo aditivo del anillo $\mathbf{Z}/p$ es un subanillo y recíprocamente.

2) Sea A un anillo y B un subanillo. Si $x \in B$, entonces $x^2 \in B$. Recíprocamente, si B es un subgrupo aditivo del anillo conmutativo A tal que 1) $x \in B \Rightarrow x^2 \in B$, 2) $2x \in B \Rightarrow x \in B$, entonces B es un subanillo.

3) Sea A un anillo. Si $a \in A$, demostrar que el conjunto de los $x \in A$ tales que $ax = xa$ es un subanillo de A.

§ 4. ANILLOS CON UNIDAD

En la definición de grupo aditivo figura como condición la existencia de un elemento cero. Del mismo modo, todo grupo multiplicativo se supone que tiene un elemento unidad. El lector ciertamente ya observó que, en la definición de anillo, la adición se comporta mejor que la multiplicación, en el sentido de que, por ejemplo, existe un elemento cero, más no se postula la existencia de un elemento unidad.

Se dice que un anillo A posee *unidad* cuando existe en A un elemento que se representará por 1, o e, o I, tal que $x1 = 1x = x$ para cualquier $x \in A$. El elemento unidad es único, siempre que exista. En efecto, si $1' \in A$ fuera tal que $x1' = 1'x = x$ para todo $x \in A$, escribiendo $x = 1'$ en $x1 = x$, y luego $x = 1$ en $1'x = x$, se obtendrá $1'1 = 1'$, $1'1 = 1$, de donde $1 = 1'$. De modo más general, si llamamos *unidad a la derecha* a todo elemento $1 \in A$ tal que $x1 = x$ y *unidad a la izquierda* a todo elemento $1'$ tal que $1'x = x$, para cualquier $x \in A$, es claro que una unidad a la derecha y una unidad a la izquierda necesariamente coinciden.

Por ejemplo, en los anillos $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$ y $\mathbf{C}$, la unidad existe y es precisamente el número 1. En el caso del anillo de las funciones reales (pág. 79), la unidad también existe y es la función constantemente igual a 1. Idem para el caso de los anillos $\mathbf{R}[x]$ y $\mathbf{C}[x]$ (págs. 80-81). En el caso del anillo $M_n(\mathbf{R})$ de las matrices reales de orden n (pág. 86) la uni-

dad es la matriz que posee todos los elementos de la diagonal principal iguales a 1 y los restantes iguales a cero. Así, en el caso $n = 2$:

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

como es fácil constatar si se calcula $a1$ y $1a$ y se obtiene los dos resultados iguales a a.

Ejemplo 1. Existen anillos sin elemento unidad. Por ejemplo, representemos por $2\mathbf{Z}$ el anillo formado por los enteros de la forma $2n$, $n \in \mathbf{Z}$, esto es los enteros pares con la adición y la multiplicación usuales. $2\mathbf{Z}$ es un anillo y además es conmutativo, pues en realidad es un subanillo de $\mathbf{Z}$ (proposición 3, pág. 90). Afirmamos que $2\mathbf{Z}$ no tiene elemento unidad. Una manera precipitada de justificar esta afirmación consistiría en sólo alegar que el elemento unidad de $\mathbf{Z}$ no pertenece a $2\mathbf{Z}$. El raciocinio completo que prueba la ausencia de un elemento unidad en $2\mathbf{Z}$ es como sigue. Supongamos que $2\mathbf{Z}$ poseyera una unidad e, esto es, que $xe = x$ para todo $x \in 2\mathbf{Z}$, donde $e \in 2\mathbf{Z}$. Entonces, $e = 2n$, $x = 2t$, donde n, $t \in \mathbf{Z}$. Luego, $4nt = 2t$ para cualquier $t \in \mathbf{Z}$. Si $t = 1$, se obtiene $2n = 1$, lo que es imposible para n entero. Luego, el anillo $2\mathbf{Z}$ no posee unidad. El mismo raciocinio puede repetirse para el anillo $p\mathbf{Z}$ de los múltiplos enteros de cualquier entero $p \geq 2$.

Ejemplo 2. Consideremos el anillo G formado por las funciones f, reales y continuas en $[a, b]$, que se anulan en a (esto es $f(a) = 0$), con relación a la adición y multiplicación definidas en la página 80 y tal que $a < b$. G es un anillo y, además, es conmutativo, pues, como el lector podrá verificar, es un subanillo del anillo F (pág. 80). Afirmamos que el anillo G no posee unidad, lo que se comprueba con el argumento siguiente: Supongamos que G posee una unidad e, tal que $fe = f$ para cualquier $f \in G$,

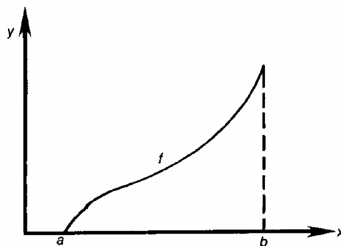


Fig. 34

donde $e \in G$. Esto quiere decir que $(fe)(x) = f(x)$, o sea que $f(x)e(x) = f(x)$ para cualesquiera $f \in G$ y $x \in [a, b]$. Consideremos, en particular, la función f definida en $[a, b]$ por $f(x) = x - a$. Es claro que esta función es continua en $[a, b]$ y que se anula en a. Luego, $f \in G$. Notemos también que $f(x) = x - a \neq 0$ siempre que $a < x \leq b$. La igualdad $f(x)e(x) = f(x)$ da, entonces, $e(x) = 1$, puesto que $a < x \leq b$. Como la función e es continua en $[a, b]$ y, en particular, en el punto a, la

igualdad $e(x) = 1$ para $a < x \leq b$ exige que $e(a) = 1$, lo que contradice $e(a) = 0$, por el hecho de ser $e \in G$. Luego, el anillo G no tiene unidad.

Consideremos un anillo A, con unidad, que con más cuidado designaremos por 1_A, y sea B un subanillo de A. Dos casos pueden presentarse:

1) La unidad 1_A pertenece a B. Entonces, es claro que B también es un anillo con unidad 1_A, la que precisamente es la unidad de A.

2) La unidad 1_A de A no pertenece a B. Es necesario, entonces, distinguir con atención dos eventualidades. Puede suceder que el anillo B no posea unidad. Es el caso, por ejemplo, del anillo **Z** y del subanillo 2**Z**: éste no sólo no contiene la unidad de **Z**, sino que tampoco tiene su propia unidad. Asimismo, puede suceder que el subanillo B no contenga 1_A, pero posea un elemento unidad $1_B \in B$ tal que $x1_B = 1_Bx = x$ para todo $x \in B$. En este caso, $1_A \neq 1_B$, pues, por hipótesis, $1_A \notin B$ y $1_B \in B$. ¿Hay en esto alguna contradicción con la unicidad del elemento unidad? No, pues las relaciones

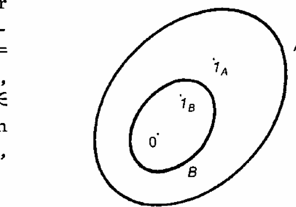


Fig. 35

$$x1_A = 1_Ax = x, \qquad x1_B = 1_Bx = x$$

se satisfacen para todo $x \in B$ (y la primera también para $x \in A$), pero 1_A no pertenece a B, lo que impide que sea $x = 1_A$ en la segunda, para concluir (después de hacer $x = 1_B$ en la primera) que $1_A = 1_B$. El ejemplo que sigue ilustra lo que se acaba de decir.

Ejemplo 3. Consideremos el anillo F de las funciones reales definidas en $[a, b]$ con $a < b$ (pág. 79) e indiquemos con H el subanillo de las funciones f reales en $[a, b]$, que se anulan en a (esto es, tales que $f(a) = 0$). Observemos, por razones de claridad, que las funciones que constituyen el subanillo H no se suponen continuas, al contrario de lo que se supuso en el ejemplo 2. El anillo F tiene unidad 1_F, a saber: la función constante 1 en $[a, b]$,

$$1_F(x) = 1 \text{ si } a \leq x \leq b.$$

Notemos que $1_F \notin H$, pues $1_F(a) = 1 \neq 0$. No obstante, el subanillo H posee su propio elemento unidad. En efecto, representemos por 1_H a la función real definida así:

$$1_H(x) = \begin{cases} 1 & \text{si } a < x \leq b \\ 0 & \text{si } x = a. \end{cases}$$

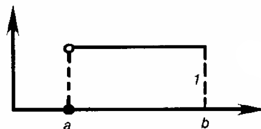


Fig. 36

Es claro que $1_H \in H$. Para probar que $f\,1_H = f$ para cualquier $f \in H$, hay que demostrar que $(f\,1_H)(x) = f(x)$, o sea $f(x)\,1_H(x) = f(x)$ para toda $f \in H$ y todo $x \in [a, b]$. Ahora bien, eso es claro si $a < x \leq b$, pues, entonces, $1_H(x) = 1$. Para $x = a$, la igualdad de arriba también es verdadera, pues $f(a) = 0$, dado que $f \in H$. Luego, 1_H es el elemento unidad del anillo H. Notemos que tal función no es continua, lo que explica la ausencia de unidad en el caso del subanillo G del ejemplo 2.

Sea A un anillo con elemento unidad, que representaremos por e para evitar confusión con el número entero 1. Si $x \in A$ y $n \in \mathbf{Z}$, las relaciones $x = ex = xe$ y la proposición 2 (pág. 90) permiten escribir $nx = (ne)x = x(ne)$. Esto muestra que, en el caso de un anillo con unidad, es posible reducir el producto nx de un entero por un elemento del anillo al producto de dos elementos ne y x del anillo, en cualquier orden. Así, en el caso de existir una unidad, una expresión tal como $ax + nx$ (donde a, $x \in A$ y $n \in \mathbf{Z}$) puede escribirse $ax + (ne)x = (a + ne)x$. Conviene advertir, empero, que en ausencia de un elemento unidad, no es correcto escribir $ax + nx = (a + n)x$, por el simple hecho de que la suma $a + n$ de un elemento del anillo con un número entero no tiene sentido.

Ejercicios

1) Sea A un anillo con unidad 1. Entonces, $(-1)x = -1x$. Si $B \subset A$ es tal que: 1) x, $y \in B \Rightarrow x + y \in B$, $xy \in B$ y 2) $-1 \in B$, entonces B es un subanillo.

2) Sea A un anillo con unidad 1. Si $1 = 0$, entonces A se reduce a un elemento. Recíprocamente, si el conjunto A consiste de un elemento a y definimos $a + a = a$, $aa = a$, entonces A resulta ser un anillo, donde a es al mismo tiempo el cero y la unidad.

3) En el anillo $\mathbf{Z}/p$, el elemento unidad es el número 1 si $p \geq 2$. Si $p = 1$, el elemento unidad es el propio 0.

4) Las matrices reales de la forma $\begin{pmatrix} x & y \\ 0 & 0 \end{pmatrix}$ forman un anillo A con relación a la adición y la multiplicación usuales entre matrices. Este anillo A no posee ninguna unidad a la derecha. A, entre tanto, tiene una infinidad de unidades a la izquierda, a saber: cada una de las matrices $\begin{pmatrix} 1 & t \\ 0 & 0 \end{pmatrix}$, $t \in \mathbf{R}$.

5) Un elemento e de un anillo A se dice *ídempotente* si $e^2 = e$. De allí resulta que $e^n = e$ $(n = 1, 2, \ldots)$. El cero es ídempotente. Si A tiene unidad, ésta es ídempotente. Recíprocamente, si $e \in A$ es ídempotente, el conjunto de los $x \in A$ tales que $xe = ex = x$ es un subanillo de A que admite a e como elemento unidad. Demostrar que, en el anillo F de las funciones reales (pág. 79), los elementos ídempotentes son, precisamente, las funciones cuyos valores son 0 ó 1. Demostrar también que, en el anillo $M_2(\mathbf{R})$ de las matrices reales de orden 2 (pág. 84), los elementos ídempotentes son las matrices

$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ tales que $a_{11}a_{22} - a_{12}a_{21} = 0$, $a_{11} + a_{22} = 1$, además de las matrices $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ y $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

6) Sea A un conjunto con dos operaciones $(x, y) \mapsto x + y$ y $(x, y) \mapsto xy$, ambas de $A \times A$ en A, tales que: 1) la adición es asociativa, 2) valen las leyes de cancelación $x + y = x + z \Rightarrow y = z$ e $y + x = z + x \Rightarrow y = z$, 3) la multiplicación es doblemente distributiva con relación a la adición, 4) existe un elemento 1 tal que $x1 = 1x = x$. Probar que la adición es conmutativa.

§ 5. HOMOMORFISMOS E ISOMORFISMOS

Dados dos anillos A y B, se llama *homomorfismo* de A en B a toda función $f : A \to B$ tal que

$$f(x + y) = f(x) + f(y), \quad f(xy) = f(x)f(y), \quad (x, y \in A),$$

o sea a toda aplicación de A en B que respete las operaciones del *anillo*. Un homomorfismo del anillo A en el anillo B es, en particular, un homomorfismo de A en B, si se consideran como grupos aditivos, según se sigue de la primera de las condiciones: $f(x + y) = f(x) + f(y)$. Todos los hechos válidos para los homomorfismos de grupos aditivos, como

$$f(0) = 0, \quad f(-x) = -f(x),$$

son, en consecuencia, válidos para los homomorfismos de anillos.

Un *isomorfismo* del anillo A en el anillo B es, por definición, cualquier homomorfismo biunívoco (o inyectivo) de A en B. Un isomorfismo entre A y B es, de acuerdo con la convención anterior (véanse las págs. 16 y 67), un isomorfismo de A sobre B. Dos anillos se dicen *isomorfos* y son considerados idénticos desde el punto de vista abstracto --esto es, de aquel en el cual lo que tiene importancia no es la naturaleza de los elementos que constituyen un anillo sino el modo en el cual ellos se combinan algebraicamente-- cuando existe al menos un isomorfismo *entre* esos anillos.

De un modo general, diremos que un anillo B es *imagen homomorfa* del anillo A cuando existe al menos un homomorfismo de A sobre B.

Como en el caso de los grupos aditivos, la función $0 : A \to B$ (pág. 13) que a cada elemento de A asocia siempre el cero de B es un homomorfismo llamado homomorfismo *cero* de A en B.

A todo homomorfismo $f : A \to A$ de un anillo A en sí mismo se le denomina *endomorfismo*. Llámase *automorfismo* de un anillo A a cualquier isomorfismo de A sobre sí mismo. La transformación idéntica $I : A \to A$ es un claro ejemplo de automorfismo.

Ejemplo 1. En el anillo **C** de los números complejos, la correspondencia $z \to \overline{z}$, que a cada número complejo le asocia su conjugado, es un automorfismo de **C**, como resulta de

$$\overline{z+w} = \overline{z} + \overline{w}, \qquad \overline{zw} = \overline{z}\,\overline{w} \qquad (z,\ w \in \mathbf{C})$$

y del hecho de que la correspondencia $z \to \overline{z}$ es biunívoca sobre **C**.

Ejemplo 2. Consideremos los anillos **Z** y $\mathbf{Z}/p$, el primero respecto a la adición y multiplicación usuales y el segundo respecto a la adición y la multiplicación módulo p. La función $r : \mathbf{Z} \to \mathbf{Z}/p$, que a cada $x \in \mathbf{Z}$ asocia el residuo $r(x)$ de la división de x entre p, es un homomorfismo, esto es,

$$r(x+y) = r(x) \overset{p}{+} r(y), \qquad r(xy) = r(x) \overset{p}{\cdot} r(y) \qquad (x,\ y \in \mathbf{Z}).$$

La primera de estas relaciones ya fue establecida antes (pág. 68). Probemos ahora la segunda. Si dividiéramos x e y entre p, tendremos

$$x = pq + m, \qquad y = pq' + m',$$

donde q y q' son los cocientes y m y m' son los residuos. Entonces, $r(x) = m$ y $r(y) = m'$. Para calcular $m \overset{p}{\cdot} m'$, dividamos mm' entre p, lo que da

$$mm' = q''p + m'',$$

96

donde q'' es el cociente y m'' el residuo. De ahí resulta que $m'' = m \overset{p}{\cdot} m'$. Ahora,

$$xy = (qp+m)(q'p+m') = (qq'p + qm' + mq')p + mm' =$$

$$= (qq'p + qm' + mq' + q'')p + m'',$$

de donde se concluye que m'' es también el residuo de la división de xy entre p

$$r(xy) = m'' = m \overset{p}{\cdot} m' = r(x) \overset{p}{\cdot} r(y),$$

como era necesario probar. Obsérvese que este homomorfismo no es un isomorfismo (o sea, no es biunívoco), pues si $p \neq 0$, como $r(p) = 0$, resulta $r(p) = r(0)$. Por otra parte, ese homomorfismo actúa sobre $\mathbf{Z}/p$, esto es el anillo $\mathbf{Z}/p$ es una imagen homomorfa del anillo **Z**.

Ejemplo 3. Se sabe, por álgebra elemental, que todo número complejo $z = x + iy$ puede representarse geométricamente por un punto (x, y) del plano de Argand-Gauss, o también por el vector que comienza en el origen y termina en ese punto. Las operaciones algebraicas entre números complejos se traducen de un modo geométrico muy sugestivo. Por ejemplo, la suma de dos números complejos corresponde a la resultante de los vectores que los representan. En lo que atañe al ejemplo que tenemos en mente, es la representación geométrica de la multiplicación la que nos interesa en especial. Si multiplicáramos un número complejo z por un número complejo $u = e^{i\varphi}$ de módulo unidad y argumento φ, el número complejo uz se representará por el punto ob-

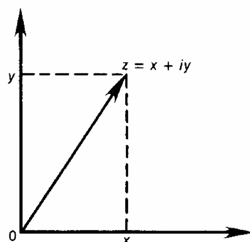


Fig. 37

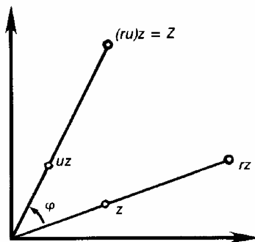


Fig. 38

tenido al hacer que el punto representativo de z gire un ángulo φ en torno del origen. Del mismo modo, si multiplicarámos z por el número real $r \geq 0$, el punto representativo de rz será obtenido a partir del punto representativo de z mediante una homotecia de centro en el origen y razón r. Escribiendo $w = ru = re^{i\varphi}$ y resumiendo, la correspondencia $z \mapsto wz$ que a cada número complejo z le asocia el número complejo $Z = wz$ es, geométricamente, la correspondencia que a cada punto del plano le asocia el punto obtenido por una rotación de ángulo φ en torno del origen, seguida de una homotecia de centro en el origen y razón r. Ahora bien, descomponiendo los números complejos z, w y Z en sus partes reales e imaginarias

$$z = x + iy, \quad w = a + ib, \quad Z = X + iY$$

y desarrollando el producto $Z = wz$, obtenemos

$$\begin{cases} X = ax - by \\ Y = bx + ay. \end{cases}$$

En virtud de las consideraciones anteriores, podemos asociar cada número complejo $w = a + ib$ con la matriz $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ de segundo orden y elementos reales que describe el sistema lineal que relaciona el punto (x, y) con su transformado (X, Y). Designemos por $f : \mathbf{C} \to M_2(\mathbf{R})$ esta correspondencia, esto es

$$f(w) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad \text{donde } w = a + ib.$$

Afirmamos que f es un isomorfismo del anillo $\mathbf{C}$ en el anillo $M_2(\mathbf{R})$. En efecto, es fácil verificar (aplicando las definiciones de suma y producto de matrices, págs. 84-85) que

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} + \begin{pmatrix} a' & -b' \\ b' & a' \end{pmatrix} = \begin{pmatrix} a + a' & -(b + b') \\ b + b' & a + a' \end{pmatrix},$$

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} a' & -b' \\ b' & a' \end{pmatrix} = \begin{pmatrix} aa' - bb' & -(ab' + ba') \\ ab' + ba' & aa' - bb' \end{pmatrix},$$

lo que prueba que $f(w) + f(w') = f(w + w')$ y que $f(w)f(w') = f(ww')$, donde $w = a + ib$ y $w' = a' + ib'$. Además, es claro que f es biunívoca. Luego, f es un isomorfismo. Notemos que f no actúa en $M_2(R)$. En efecto, una matriz $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ es de la forma $f(w) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ si, y sólo si, $a_{11} = a_{22}$ y $a_{12} = -a_{21}$. Por lo tanto, existen matrices como $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, que no pertenecen al conjunto de los valores $f(C)$ de f. La colección de las matrices reales de la forma $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ o, de modo equivalente, de las matrices reales $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ tales que $a_{11} = a_{22}$ y $a_{12} = -a_{21}$, constituye un anillo con respecto a la adición y multiplicación de matrices y, en realidad, un subanillo de $M_2(R)$ isomorfo al anillo C de los números complejos.

Una observación oportuna es la siguiente: en el caso de los grupos aditivos y multiplicativos, la matemática elemental ofrece una gran variedad de ejemplos alusivos y naturales de homomorfismos e isomorfismos. No sucede lo mismo en el caso de los anillos, en especial en lo que se refiere a homomorfismo. Si bien no hemos definido aún el muy importante concepto de ideal, no cuesta nada adelantar que esta ausencia de ejemplos naturales se debe a la circunstancia de que en los anillos Q, R, C, $M_2(R)$ y $M_2(C)$ casi no existen ideales. No se debe inferir de ello la falsa idea de que la noción de homomorfismo entre anillos no presenta gran interés: lo correcto es que la mayor parte de los ejemplos importantes de homomorfismos de anillo corresponde más al espíritu del álgebra moderna que al de la matemática elemental.

A título de ejemplo de lo que vimos en el caso de grupos (pág. 71), se llama *antihomomorfismo* de un anillo A en otro anillo B a toda aplicación $f : A \to B$ tal que

$$f(x + y) = f(x) + f(y), \quad f(xy) = f(y)f(x) \quad (x, y \in A)$$

(nótese la inversión del orden en el segundo miembro de la última ecuación). Las nociones de antiisomorfismo, antiendomorfismo y antiautomorfismo se definen de manera similar. Es claro que si uno de los dos anillos A y B fuera conmutativo, las nociones de homomorfismo y antihomomorfismo de A en B coinciden.

Ejemplo 4. Consideremos, en el anillo $M_2(R)$ (pág. 84), la aplicación $a \mapsto {}^t a$ que a cada matriz $a = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ le asocia su transpuesta

${}^t a = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}$, que se obtiene escribiendo la primera fila de a como primera columna de ${}^t a$ y la segunda fila de a como segunda columna de ${}^t a$. A partir de las definiciones, se verifica fácilmente que

$$ {}^t(a+b) = {}^t a + {}^t b, \qquad {}^t(ab) = {}^t b\, {}^t a, $$

para cualesquiera que sean a, $b \in \mathbf{M}_2(\mathbf{R})$. Como la aplicación $a \mapsto {}^t a$ es obviamente biunívoca y sobre $\mathbf{M}_2(\mathbf{R})$, vemos que la operación de transposición es un antiautomorfismo de $\mathbf{M}_2(\mathbf{R})$.

Ejercicios

1) Sea F el anillo de las funciones reales de variable real definidas en un intervalo $[a, b]$ (pág. 79). Fijado un x tal que $a \le x \le b$, mostrar que la función $\Phi : F \to \mathbf{R}$ definida por $\Phi(f) = f(x)$ es un homomorfismo del primer anillo sobre el segundo, mas no es un isomorfismo.

2) Sea $f : A \to B$ un homomorfismo del anillo A en el anillo B. Para todo subanillo $X \subset A$, $f(X)$ es un subanillo de B. En particular, $f(A)$ es un subanillo de B. Para todo subanillo $Y \subset B$, $f^{-1}(Y)$ es un subanillo de A. Si A fuera conmutativo, el anillo $f(A)$ también será conmutativo.

3) Sea A un anillo. Consideremos el producto cartesiano $\mathbf{Z} \times A$ y definamos en el mismo una adición y una multiplicación mediante

$$ (m, x) + (n, y) = (m + n, x + y) $$

$$ (m, x)(n, y) = (mn, my + nx + xy) $$

Demostrar que $\mathbf{Z} \times A$ es un anillo con respecto a tales operaciones y que $(1, 0)$ es su unidad. Mostrar también que la función $x \mapsto (0, x)$ de A en $\mathbf{Z} \times A$ es un isomorfismo. En particular, todo anillo es isomorfo a un subanillo de otro anillo con unidad.

4) Sea G un grupo aditivo. Demostrar que el conjunto $E(G)$ de los endomorfismos del grupo aditivo G forma un anillo con unidad del siguiente modo: a) si f y g fueran endomorfismos de G, entonces $f + g$ es el endomorfismo de G tal que $(f + g)(x) = f(x) + g(x)$ para cualquier $x \in G$ y fg es el endomorfismo de G definido por $(fg)(x) = f[g(x)]$, cualquiera que sea $x \in G$; b) la unidad de $E(G)$ es la transformación idéntica de G en sí mismo.

5) Sea A un anillo con unidad (el cual, por tanto, también es un grupo aditivo). Todo elemento $a \in A$ determina un endomorfismo $m_a : A \to A$ del grupo aditivo A, a saber: la función dada por $x \mapsto ax$. Demostrar que la transformación que a cada $a \in A$ le asocia $m_a \in E(A)$ es un isomorfismo del anillo A en el anillo $E(A)$ (véase el ejercicio anterior para la definición del anillo de endomorfismos de un grupo aditivo). En particular, todo anillo con unidad es isomorfo a un subanillo del anillo con unidad de los endomorfismos de un grupo aditivo *(Teorema de Cayley)*. A partir del ejercicio 3, mostrar que todo anillo (con o sin unidad) es isomorfo a un subanillo del anillo con unidad de los endomorfismos de

un grupo aditivo. Si en la construcción precedente sustituyéramos $x \mapsto ax$ por $x \mapsto xa$, obtendríamos un antiisomorfismo en vez de un isomorfismo.

§ 6. CUERPOS CONMUTATIVOS

Entre los anillos con unidad que ocurren naturalmente en matemática elemental figuran los cuerpos, que se destacan por su simplicidad algebraica desde el punto de vista de la divisibilidad. Antes de entrar a definir los cuerpos necesitamos puntualizar el concepto de inverso de un elemento en un anillo con unidad.

Sea A un anillo con unidad 1. Diremos que un elemento $x \in A$ es *inversible* en A cuando exista en A un elemento (indicado por x^{-1}), denominado el *inverso* de x en A, tal que

$$xx^{-1} = x^{-1}x = 1.$$

Un tal elemento x^{-1} es necesariamente único, como se constata fácilmente gracias al raciocinio empleado en la demostración de la unicidad del inverso en grupos multiplicativos (pág. 54).

Advirtamos que la unidad 1 de A es inversible en A y coincide con su inverso en A. Por otro lado, el cero 0 de A no es inversible en A si es que $1 \neq 0$.

En un anillo A con unidad, si $x, y \in A$ y y^{-1} existe en A, entonces x e y conmutan (esto es $xy = yx$) si, y sólo si, x e y^{-1} conmutan también (esto es $xy^{-1} = y^{-1}x$). En efecto, de $xy = yx$ se deduce que $y^{-1}xy\, y^{-1} = y^{-1}yx\, y^{-1}$, de donde $y^{-1}x = xy^{-1}$ y, recíprocamente, de esta última resulta la primera por un raciocinio análogo. Gracias a tales hipótesis de conmutatividad, se define el *cociente* de x entre y, o la *fracción de numerador* x y *denominador* y, mediante

$$\frac{x}{y} = xy^{-1} = y^{-1}x,$$

y se usa también la notación x/y para representarla.

Proposición 1. En todo anillo A con unidad se tiene:

1) Si $x \in A$ es inversible en A, entonces $x^{-1} \in A$ es inversible en A y $(x^{-1})^{-1} = x$.

2) Si $x, y \in A$ son inversibles en A, entonces xy es inversible en A y $(xy)^{-1} = y^{-1}x^{-1}$.

Demostración. Basta repetir el raciocinio hecho en el caso de la proposición 1, páginas 54-55, puntos 1 y 2. QED

Un cuerpo K es un anillo con unidad $1 \neq 0$, donde todo elemento diferente de 0 es inversible en K.

Ejemplo 1. Los cuerpos conmutativos más importantes en matemática elemental son el cuerpo **Q** de los números racionales, el cuerpo **R** de los números reales y el cuerpo **C** de los números complejos.

Proposición 2. En todo cuerpo K se tiene:

1) Si x, $y \in A$ y $xy = 0$, entonces $x = 0$ o $y = 0$.

2) Si x, y, $z \in A$ y $x \neq 0$, entonces $xy = xz \Rightarrow y = z$ y $yx = zx \Rightarrow y = z$.

Demostración. Si x, $y \in A$ y $xy = 0$, entonces o bien $x = 0$ o $x \neq 0$. En este último caso, x es inversible en A y, entonces, $xy = 0$ implica $x^{-1}xy = 0$, de donde $y = 0$.

Por otro lado, si x, y, $z \in A$ y $x \neq 0$, entonces x es inversible en A, de modo que $xy = xz$ implica $x^{-1}xy = x^{-1}xz$, de donde $y = z$. La propiedad que acaba de establecerse se denomina *ley de cancelación a la izquierda*. Análogamente se establece la *ley de cancelación a la derecha*, que se expresa mediante la segunda parte del ítem 2. QED

Ejemplo 2. Dado el número entero $p \geq 1$, consideremos el anillo $\mathbf{Z}/p$, que es conmutativo y tiene elemento unidad (pág. 82). Procuremos descubrir los valores de p para los cuales $\mathbf{Z}/p$ es un cuerpo. En primer lugar, para que la unidad de $\mathbf{Z}/p$ sea diferente de su cero es obviamente necesario y suficiente que $p \geq 2$. En el caso de que p no sea primo, tendrá una factorización $p = xy$ en el sentido de la multiplicación usual, siendo x e y números enteros tales que $1 < x < p$, $1 < y < p$. Entonces x, $y \in \mathbf{Z}/p$ y tenemos que $x \overset{p}{\cdot} y = 0$. Como $x \neq 0$ e $y \neq 0$, vemos, por el ítem 1 de la proposición 2, que $\mathbf{Z}/p$ no puede ser un cuerpo. Por otra parte, supongamos ahora que $p \geq 2$ sea primo. Vamos a emplear el siguiente teorema de aritmética elemental: si x e y fueran dos enteros y $d \geq 0$ designa a su máximo común divisor, entonces existen dos números enteros u y v tales que $ux + vy = d$. En particular, si x e y fueran primos entre sí (o sea si $d = 1$), existen dos enteros u y v tales que $ux + vy = 1$. Ahora, dado $x \in \mathbf{Z}/p$ con $x \neq 0$, como p es primo, resulta que x y p son primos entre sí. Luego, existen dos números enteros u y v tales que $ux + vp = 1$. Si se divide u entre p, llamando q al cociente y r al residuo, se tendrá $u = qp + r$, de donde $(qp + r)x + vp = 1$, o sea, $rx = sp + 1$, donde $s = -(v + qx)$. Ahora, $r \in \mathbf{Z}/p$ y la igualdad $rx = sp + 1$ muestra que $r \overset{p}{\cdot} x = 1$. Luego x es inversible en $\mathbf{Z}/p$ y r es precisamente el inverso de x en $\mathbf{Z}/p$. Si esto se cumple para todo $x \neq 0$ en $\mathbf{Z}/p$, vemos que $\mathbf{Z}/p$ es un cuerpo. En resumen, $\mathbf{Z}/p$ es un cuerpo si, y sólo si, $p \geq 2$ es primo.

Observación. Los cuatro ejemplos de cuerpos dados arriba son todos conmutativos. Existen ejemplos importantes de cuerpos no conmutativos, como el cuerpo de los cuaterniones, cuyo análisis (si bien sencillo) escapa los alcances de la presente monografía. (Véase más adelante el ejercicio 9.) Los cuerpos conmutativos son con frecuencia llamados *campos*, en tanto que los cuerpos arbitrarios, conmutativos o no, se conocen a veces como *anillos de división*.

Un *subcuerpo* L de un cuerpo K es un subanillo L de K que es un cuerpo. Notemos que la unidad de K es también unidad de L. En efecto si

se indican con 1_K y 1_L las unidades de K y L, respectivamente, se tendrá $1_K 1_L = 1_L$, pues 1_K es la unidad de K, así como $1_L 1_L = 1_L$ (pues 1_L es la unidad de L), lo que implica $1_K 1_L = 1_L 1_L$, lo que implica a su vez $1_K = 1_L$, por el ítem 2 de la proposición 2. Basta, pues, indicar con 1 a la unidad de K y de L. Observemos también que todo $x \neq 0$ en L tiene el mismo inverso tanto en K como en L. En efecto, indicando con x_L^{-1} el inverso de x en L, vemos que

$$xx_L^{-1} = x_L^{-1}x = 1$$

y, dada la unicidad del inverso, concluimos que x_L^{-1} es el inverso x_K^{-1} de x en K. Es posible, pues, indicar simplemente por x^{-1} el inverso de x tanto en K como en L. Por ejemplo, **Q** es un subcuerpo de **R** tanto como de **C**, por ser **R** un subcuerpo de **C**.

La proposición 2 expresa propiedades sencillas e importantes de los cuerpos, pero éstas constituyen privilegio exclusivo de los mismos. En realidad, dichas propiedades son equivalentes entre sí (como resulta de la proposición 3) y en el caso de los anillos conmutativos sirven de punto de partida para la definición del importante concepto de dominio de integridad. Este es un ejemplo de una actitud frecuente en matemática, a ser usada juiciosamente, que consiste en usar propiedades importantes mas no características de ciertos sistemas (los cuerpos en el caso presente) para definir otros sistemas (los dominios de integridad).

Proposición 3. Para todo anillo A, las siguientes propiedades son equivalentes:

1. Si x, $y \in A$ y $xy = 0$, entonces $x = 0$ o $y = 0$.

2i. Si x, y, $z \in A$ y $x \neq 0$, entonces $xy = xz \Rightarrow y = z$.

2d. Si x, y, $z \in A$ y $x \neq 0$, entonces $yx = zx \Rightarrow y = z$.

Demostración. Supongamos verdadera la propiedad del punto 1. Entonces, en el caso del punto 2i, podemos escribir $xy = xz$ en la forma $x(y - z) = 0$ y, como $x \neq 0$, concluimos que $y - z = 0$, de donde $y = z$. Recíprocamente, admitamos la propiedad del punto 2i. Entonces, en el caso del punto 1, o bien $x = 0$ o $x \neq 0$ y, escribiendo $xy = 0$ en la forma $xy = x0$, resulta $y = 0$. Eso prueba la equivalencia entre las condiciones 1 y 2i de la proposición. Se procede de modo análogo para 1 y 2d. QED

Un *dominio de integridad* es un anillo conmutativo que posee las tres propiedades equivalentes 1, 2i y 2d de la proposición 3, las cuales se denominan *ley de cancelación*, *ley de cancelación a la izquierda* y *ley de cancelación a la derecha*. Por la proposición 2, todo cuerpo conmutativo (o campo) es un dominio de integridad; y, de modo más general, todo subanillo de un cuerpo conmutativo es un dominio de integridad. Por consiguiente, el motivo por el cual un dominio de integridad se supone conmutativo es por ser posible construir su cuerpo de fracciones, siendo el anillo dado isomorfo a un subanillo de tal cuerpo, que es conmutativo (véase más adelante el ejercicio 8).

Ejemplo 3. El dominio de integridad más importante en matemática elemental es **Z**, que no es un cuerpo, pues (salvo por 1 y - 1) ningún elemento de **Z** es inversible en **Z**, aunque lo sea en **Q**.

Ejemplo 4. Consideremos el anillo conmutativo R$[x]$ de los polinomios reales de una variable real x (págs. 80-81). Afirmamos que R$[x]$ es un dominio de integridad. En efecto, sean

$$p(x) = a_0 + a_1x + \dots + a_mx^m, \quad q(x) = b_0 + b_1x + \dots + b_nx^n$$

dos polinomios pertenecientes a R$[x]$ y mostremos que si $p \neq 0$ y $q \neq 0$, entonces $pq \neq 0$. Ahora bien, por el principio de identidad de polinomios, si $p \neq 0$, entonces algún coeficiente de p debe ser diferente de cero. Sea i el menor entero tal que $0 \leq i \leq m$, $a_i \neq 0$. Análogamente, sea j el menor entero tal que $0 \leq j \leq n$, $b_j \neq 0$. Si nos remitimos a la fórmula que define el producto pq (pág. 81), vemos que el coeficiente de x^{i+j} es precisamente $a_ib_j \neq 0$, luego $pq \neq 0$, como queríamos. Por otra parte, obsérvese que R$[x]$ no es un cuerpo. En efecto, si $p \neq 0$ y $q \neq 0$ fueran polinomios de grados m y n, respectivamente, entonces $a_m \neq 0$ y $b_n \neq 0$. Por tanto, es claro que $a_mb_n \neq 0$, de donde resulta que pq es de grado $m + n$. Una consecuencia de tal observación es que todo polinomio $p \neq 0$ de grado $m \geq 1$ es no inversible en R$[x]$. En efecto, admitiendo que p sea inversible y que q sea su inverso, entonces $q \neq 0$ y q debe tener grado $n \geq 0$, de donde resultaría que pq es de grado $m + n \geq m \geq 1$, lo que es absurdo, pues $pq = 1$ es de grado 0. Todo lo que se acaba de indicar para R$[x]$ se repite sin alteración para C$[x]$ (pág. 81).

Ejercicios

1) Para que un anillo K sea un cuerpo es necesario y suficiente que el conjunto K^* de los elementos diferentes de cero en K sea un grupo con respecto a la operación de multiplicación de K.

2) Para que un subanillo L de un cuerpo K sea un subcuerpo de K es necesario y suficiente que L contenga el inverso en K de todo elemento $x \in L$, $x \neq 0$ y que L contenga a la unidad.

3) Si $f : K \to A$ fuera un homomorfismo de anillo del cuerpo K en el anillo A, entonces o f es un isomorfismo y $f(K)$ es un subanillo de A que es un cuerpo, o $f(K)$ se reduce al cero de A.

4) En un anillo K con unidad $1 \neq 0$, si todo elemento $x \in K$ diferente de cero tiene un inverso a la derecha en K, esto es si existe un elemento $y \in K$ tal que $xy = 1$, entonces K es un cuerpo. Se procede de modo análogo para inversos a la izquierda.

5) El anillo conmutativo F de las funciones reales definidas en un intervalo de **R** (pág. 79) no es un anillo de integridad si $a \neq b$.

6) Si A es un subanillo no reducido al cero del cuerpo conmutativo K y A^* es el conjunto no vacío de los elementos de A diferentes de cero, el conjunto de las fracciones x/y, donde $x \in A$ e $y \in A^*$, es un subcuerpo de K que contiene a A, que está contenido en cualquier subcuerpo de

K que contenga a A. (Cuando todo elemento de K es una fracción x/y, donde $x \in A$ y $y \in A^*$, se dice que K es un cuerpo de fracciones de A).

7) Si K fuera un cuerpo conmutativo y $t, x, y, z \in K$, con $y \neq 0$, $z \neq 0$, entonces

$$t/y = x/z \Leftrightarrow tz = xy,$$

$$t/y + x/z = (tz + xy)/yz,$$

$$t/y \cdot x/z = (tx)/(yz).$$

8) Sea A un dominio de integridad no reducido al cero e indiquemos con A^* el conjunto no vacío de los elementos de A diferentes de cero. En el producto cartesiano $A \times A^*$ introduzcamos la relación de equivalencia que consiste en escribir $(t, y) \sim (x, z)$ si $tz = xy$. Sea K el conjunto cociente de $A \times A^*$ con respecto a tal relación de equivalencia. Probar que K es un cuerpo conmutativo cuando definimos la suma y el producto de la clase de equivalencia de (t, y) con la clase de equivalencia de (x, z) como las clases de equivalencia de $(tz + xy, yz)$ y (tx, yz), respectivamente. Probar que la correspondencia f que a cada $t \in A$ le asocia la clase de equivalencia $f(t) \in K$ de (ty, y), donde $y \in A^*$, es un isomorfismo de anillo de A en K y que K es un cuerpo de fracciones de su subanillo $f(K)$ en el sentido del ejercicio 5. (Tomar el ejercicio 6 como modelo para la construcción hecha en este ejercicio).

9) El subconjunto H de $M_2(\mathbf{C})$ de las matrices complejas de dos filas y dos columnas de la forma

$$q = \begin{pmatrix} z & w \\ \overline{w} & \overline{z} \end{pmatrix},$$

donde $z, w \in \mathbf{C}$ es un subanillo no conmutativo que contiene a la matriz unidad de $M_2(\mathbf{C})$. Demostrar que H es un cuerpo, pues todo elemento $q \neq 0$ tiene un inverso en H dado por

$$q^{-1} = \begin{pmatrix} \overline{z}/\Delta & -w/\Delta \\ \overline{w}/\Delta & -z/\Delta \end{pmatrix},$$

donde $\Delta = |z|^2 + |w|^2 > 0$. Cada elemento de H se denomina *cuaternión de Hamilton).*

4

ESPACIOS VECTORIALES Y ALGEBRA LINEAL

La noción de vector tiene su origen en la mecánica, a través de los conceptos de desplazamiento, velocidad, aceleración y fuerza, así como en la geometría, en la noción de segmento orientado. Los vectores pueden estar sujetos a dos operaciones básicas: la adición de dos vectores y la multiplicación de un escalar (número) por un vector, los cuales tienen un vector por resultado. Con el advenimiento de la geometría analítica, el concepto de vector pasó a ser expresado mediante su descomposición con respecto a los ejes de un sistema de coordenadas dado; y, según la concepción de Descartes, cada plano con un sistema de coordenadas fijo pasó a ser identificado naturalmente con el cuadrado cartesiano $\mathbf{R}^2$, producto cartesiano del conjunto $\mathbf{R}$ de los números reales por sí mismo. Así, $\mathbf{R}^2$ es un ejemplo elemental de espacio vectorial; de modo análogo, tenemos el espacio vectorial tridimensional $\mathbf{R}^3$ en que vivimos. Generalizando, tenemos el espacio vectorial $\mathbf{R}^n$, potencia cartesiana de $\mathbf{R}$ tomado como factor n veces, donde $n = 1$, 2, 3, 4, ... $\mathbf{R}^n$ es un espacio vectorial de dimensión n --en un cierto sentido $\mathbf{R}^n$ es el espacio vectorial real de dimensión n más general, salvo isomorfismos.

Por otra parte, existen buenas razones para considerar espacios vectoriales reales definidos independientemente de sistemas de coordenadas, esto es definidos axiomática o intrínsecamente. Entre ellos se encuentran naturalmente espacios de dimensión finita o infinita. Finalmente, es útil considerar espacios vectoriales sobre un cuerpo K cualquiera, en vez de sólo espacios vectoriales reales, es decir sobre el cuerpo $\mathbf{R}$. Según los puntos de vista usuales en matemática y sus aplicaciones, los dos casos más relevantes son los de los espacios vectoriales reales ($K = \mathbf{R}$) y los espacios vectoriales complejos ($K = \mathbf{C}$).

El estudio de los espacios vectoriales es lo que se denomina álgebra lineal, un capítulo de la matemática bastante más joven que el cálculo infinitesimal, pero hoy en día tan importante como éste en las aplicaciones corrientes, a tal punto que las dos puertas de acceso a la matemática universitaria son precisamente el cálculo infinitesimal y el álgebra lineal, a las cuales se les une una tercera parte igualmente importante: la computación. Estos son algunos de los factores que llevaron al concepto de vector como base del álgebra lineal, que se amplía con la noción de tensor, base del álgebra multilineal.

§ 1. ESPACIOS VECTORIALES

En matemática elemental se encuentran muchos ejemplos de conjuntos cuyos elementos pueden sumarse y también multiplicarse por escalares. Entre ellos mencionaremos los siguientes:

Ejemplo 1. Un desplazamiento de un punto en un plano P (y, de modo similar, en el espacio tridimensional en que vivimos) desde A hasta B depende de tres datos: 1) la distancia que separa a A y B; 2) si $A \neq B$, la dirección de la recta que pasa por A y B; 3) si $A \neq B$, el sentido en que se va de A a B en tal recta. En el caso particular de $A = B$, el desplazamiento en cuestión queda determinado tan sólo por la distancia 0 que separa a A y B; la dirección y el sentido carecen de significado. De ahí la representación por el segmento orientado $\overrightarrow{AB}$, con una flecha que indica *de A para B*. Entretanto, es conveniente referir todo esto a un origen fijo, o sea a un punto 0 del plano P con respecto al cual representamos el desplazamiento en cuestión a través de un segmento orientado x, llamado *vector*, de origen 0 y de la misma longitud que $\overrightarrow{AB}$ y, si $A \neq B$, con la misma dirección y el mismo sentido que $\overrightarrow{AB}$; si $A = B$ el vector x empieza y termina en 0. Se acostumbra escribir $B = A + x$ para indicar que B se obtiene de A mediante el desplazamiento x. Si sometemos ahora al punto B a un desplazamiento hasta el punto C, representamos a $\overrightarrow{BC}$ con respecto al origen 0 por el vector y, de la misma longitud y con la misma dirección y sentido en el caso en que $B \neq C$; si $B = C$, entonces y comienza y termina en 0. El desplazamiento total, de A hasta B, más el de B hasta C, o sea $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$ (como se escribe simbólicamente, cuando se lo refiere al origen 0) será la diagonal del paralelogramo *(regla del paralelogramo)* construido a partir de x e y como lados, vector que representamos por $x + y$, como es natural. No vamos a enumerar los casos simples que requieren interpretaciones sencillas por separado, cuando $A = B$, o $B = C$, o $A \neq B$ y $B \neq C$, pero tales que $\overrightarrow{AB}$ y $\overrightarrow{BC}$ tengan la misma dirección. Por otra parte, consideremos dos puntos distintos A y B y un desplazamiento desde A hasta el punto B', en la dirección y sentido de A hacia B; llamemos $\lambda' = AB'/AB$ al cociente de las distancias respectivas. Cuando el desplazamiento $\overrightarrow{AB'}$ se refiere al origen 0 (y visto que $AB' = \lambda' \cdot AB$) se obtiene un vector que naturalmente se representa por $\lambda' x$. Entretanto, sean nuevamente A y B dos puntos distintos y consideremos un desplazamiento en su dirección, pero en el sentido opuesto al de A hacia B, desde A hasta el punto B'', y escribamos $\lambda'' = AB''/AB$ para el cociente de las distancias respectivas. Refiriendo el desplazamiento $\overrightarrow{AB''}$ al origen 0 y una vez que $AB'' = \lambda'' \cdot AB$, se obtiene un vector representado naturalmente por $(-\lambda'')x = -\lambda'' x$, donde el signo negativo tiene la finalidad de indicar que el desplazamiento fue tomado en el sentido opuesto. Si $A = B$, de modo que x comienza y termina en 0, el producto λx se define como el vector que comienza y termina en 0, cualquiera que sea el número real λ. Se constata que el plano P se convierte así en un espacio vectorial, históricamente quizá el primer ejemplo de tal concepto.

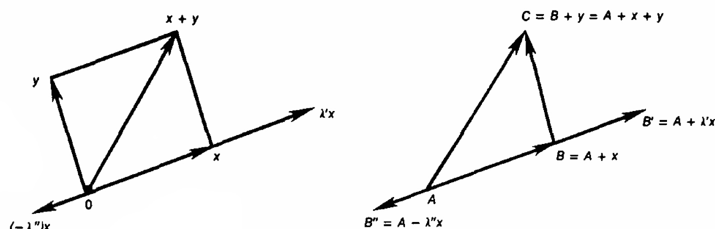


Fig. 39

Ejemplo 2. Con el advenimiento de la geometría analítica en el plano $\mathbf{R}^2$ (y, de modo similar, en el espacio tridimensional $\mathbf{R}^3$), cada punto del plano $\mathbf{R}^2$ pasó a ser expresado por el par ordenado de sus coordenadas. De igual modo, cada vector x de $\mathbf{R}^2$ pasó a ser identificado, a partir del origen 0, con el par ordenado (x_1, x_2) de las coordenadas de su extremo final, escribiéndose entonces $x = (x_1, x_2)$. Con esta notación, la adición de dos vectores $x = (x_1, x_2)$ e $y = (y_1, y_2)$ de $\mathbf{R}^2$ pasa a ser descrita por $x + y = (x_1 + y_1,\ x_2 + y_2)$ y la multiplicación de un número real λ por un vector $x = (x_1, x_2)$ se vuelve $\lambda x = (\lambda x_1, \lambda x_2)$. Tal concepción vectorial de $\mathbf{R}^2$, como la análoga de $\mathbf{R}^3$, para citar el caso que visualizamos intuitivamente, es una combinación feliz de los puntos de vista de la mecánica y la geometría (véase el ejemplo 1) y de Descartes con su geometría analítica, que nos lleva a una concepción vectorial de $\mathbf{R}^n$ ($n = 1, 2, 3, 4, \ldots$), la cual ya no visualizamos, pero que es muy fecunda por su contenido álgebro-analítico-geométrico, y que conduce al álgebra lineal en dimensión finita.

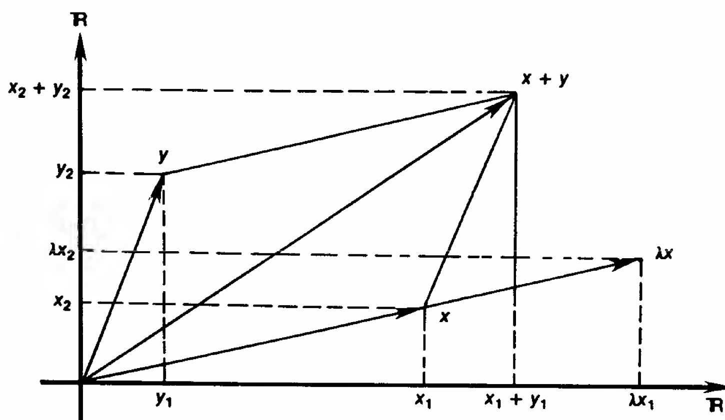


Fig. 40

Ejemplo 3. Consideremos el conjunto F de las funciones reales de variable real definidas en un intervalo $[a, b]$. Si $f, g \in F$, introduzcamos la suma $f + g \in F$ como en el ejemplo 3, § 1, capítulo 2. Añadamos ahora la definición del producto $\lambda f \in F$, donde $\lambda \in \mathbf{R}$ y $f \in F$, como la función que satisface

$$(\lambda f)(x) = \lambda \cdot f(x)$$

para $x \in [a, b]$, o sea, λf es la función que en el punto x toma el valor $\lambda \cdot f(x)$. Tenemos ahora dos operaciones en F, una de adición y otra de multiplicación (pero, en este caso, la multiplicación no se trata de la considerada en el ejemplo 2, § 1, capítulo 3). Entonces F será un nuevo ejemplo de espacio vectorial de dimensión infinita, suponiendo que $a < b$. El hecho de que las funciones reales que constituyen F estén definidas solamente en un intervalo $[a, b]$ es una restricción que puede ser suprimida, ya que también se obtiene un espacio vectorial mediante consideraciones semejantes, hechas en el conjunto F de las funciones

reales definidas en un conjunto E. Además, si $E = \{1, \ldots, n\}$ fuera el conjunto de los enteros de 1 a n, entonces $F = \mathbf{R}^n$.

El examen de los ejemplos dados, como también de otros, lleva a sintetizarlos mediante el concepto de espacio vectorial.

Un *espacio vectorial sobre un cuerpo* es un conjunto E al cual está asociado un cuerpo K, en el que están dadas dos operaciones, una de adición y otra de multiplicación que satisfacen las siguientes condiciones:

1. Con respecto a la adición, E es un grupo aditivo (se satisfacen, pues, las condiciones 1 al 5 de las págs. 47 y 48).

2. La operación de multiplicación es una función definida en $K \times E$, con valores en E que, a cada par ordenado (λ, x), donde $\lambda \in K$ y $x \in E$, asocia un elemento $\lambda x \in E$, llamado producto de los factores λ y x.

3. La multiplicación es asociativa:

$$(\lambda\mu)x = \lambda(\mu x) \quad (\lambda, \mu \in K, \ x \in E).$$

4. La unidad de K actúa como unidad de E:

$$1x = x \quad (x \in E).$$

5. La multiplicación es doblemente distributiva con relación a la adición:

$$\lambda(x + y) = \lambda x + \lambda y, \quad (\lambda + \mu)x = \lambda x + \mu x \quad (\lambda, \mu \in K, \ x, y \in E).$$

Los elementos del espacio vectorial E se denominan *vectores* en tanto que los elementos del cuerpo K se llaman *escalares*. El cero, 0, de E se llama origen de E.

Los ejemplos considerados anteriormente ilustran el concepto de espacio vectorial sobre el cuerpo $\mathbf{R}$. En el tercer ejemplo basta sustituir $\mathbf{R}$ por K para tener un espacio vectorial sobre K. Todo cuerpo K es un espacio vectorial sobre K con respecto a la adición y a la multiplicación en K. Un espacio vectorial sobre $\mathbf{R}$ se llama *espacio vectorial real*. Un espacio vectorial sobre $\mathbf{C}$ se llama *espacio vectorial complejo*. Son los dos tipos de espacios vectoriales más comúnmente empleados en análisis y geometría. Entretanto, en estas mismas dos disciplinas, así como en álgebra hay razones para abordar espacios vectoriales sobre otros cuerpos, como el cuerpo de los cuaterniones de Hamilton, o los llamados cuerpos p-ádicos.

Proposición 1. En todo espacio vectorial E sobre un cuerpo K se tiene:

1. $0x = \lambda 0 = 0$,
2. $-\lambda x = (-\lambda)x = \lambda(-x)$ y $(-1)\,x = -x$,
3. $(-\lambda)(-x) = \lambda x$,
4. $\lambda(x - y) = \lambda x - \lambda y$, $\quad (\lambda - \mu)\,x = \lambda x - \mu x$.

La demostración es análoga a la de la proposición 1, § 2, capítulo 3 y será omitida.

Ejercicios

1) Mostrar que en un espacio vectorial E sobre un cuerpo K, $\lambda x = 0$ ($\lambda \in K$, $x \in E$) implica $\lambda = 0$ o $x = 0$.

2) Consideremos un espacio vectorial E sobre un cuerpo K. Además de la suma $x + y$ y del producto λx ($x, y \in E$, $\lambda \in K$), ya definidos, introduzcamos un nuevo producto definido por $\lambda \circ x = 0$ ($\lambda \in K$, $x \in E$). Mostrar que E y K satisfacen todos los axiomas de la definición de espacio vectorial con respecto a la suma $x + y$ ya dada y al nuevo producto $\lambda \bullet x = 0$, excepto por $1 \bullet x = x$ cuando $E \neq 0$ (esto es, cuando E no se reduce al origen). Luego, en la definición de espacio vectorial E sobre un cuerpo K, el axioma $1x = x$ no es consecuencia de los demás.

3) Probar que en un espacio vectorial E sobre un cuerpo K, el axioma de la conmutatividad de la adición es consecuencia de los demás, cuando se calcula $(1 + 1)(x + y)$ de dos modos posibles.

4) Consideremos el cuerpo $K = Z/p$ (de p elementos) de los enteros módulo p (véase el ejemplo 2, §6, capítulo 3) y el espacio vectorial $E = K^n$ (de p^n elementos) sobre el cuerpo K, donde $p \geq 2$ es un entero primo y $n = 1, 2, 3, \ldots$. Escribir la tabla de multiplicación de escalares por vectores para valores pequeños de $p = 2, 3, 5, \ldots$ y $n = 1, 2, 3, \ldots$.

5) Si E es un espacio vectorial sobre un cuerpo K, entonces E será también un espacio vectorial sobre un subcuerpo L de K con respecto a la misma adición en E y a la restricción a L de la multiplicación entre K y E. En particular, todo espacio vectorial complejo es, automáticamente, un espacio vectorial real, llamado espacio vectorial real subyacente al espacio vectorial complejo dado.

6) Sea E un espacio vectorial real. Construyamos un espacio vectorial complejo $E_C = E \times E = E^2$, llamado la complejificación de E, de la siguiente manera. La adición en E_C se define de la manera usual: si $x_1, x_2, y_1, y_2 \in E$ y $x = (x_1, x_2)$ e $y = (y_1, y_2) \in E_C$, entonces

$$x + y = (x_1 + y_1,\ x_2 + y_2) \in E_C.$$

La multiplicación entre $\mathbf{C}$ y E_C se define de la siguiente manera: si $\lambda_1, \lambda_2 \in \mathbf{R}$, $x_1, x_2 \in E$ y $\lambda = \lambda_1 + i\lambda_2 \in \mathbf{C}$, $x = (x_1, x_2)$, entonces $\lambda x = (\lambda_1 x_1 - \lambda_2 x_2,\ \lambda_1 x_2 + \lambda_2 x_1) \in E_C$. Probar que E_C es un espacio vectorial complejo. Adviértase que la construcción de E_C a partir de E es una generalización de la construcción de $\mathbf{C}$ a partir de $\mathbf{R}$.

7) Probar que el producto cartesiano $E = E_1 \times \ldots \times E_n$ de los espacios vectoriales $E_1, \ldots, E_n$ sobre un cuerpo K es, naturalmente, un espacio vectorial sobre K del modo siguiente. Si $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in E$, donde $x_i, y_i \in E_i$ ($i = 1, \ldots, n$), $\lambda \in K$, definamos $x + y = (x_1 + y_1, \ldots, x_n + y_n) \in E$, $\lambda x = (\lambda x_1, \ldots, \lambda x_n) \in E$. En particular, la potencia cartesiana n-ésima K^n es un espacio vectorial sobre K. Análogamente, probar que el producto cartesiano $E = E_1 \times \ldots \times E_n \times \ldots$ de los espacios vectoriales $E_1, \ldots, E_n, \ldots$ sobre el cuerpo K es, de modo natural, un espacio vectorial sobre K del modo siguiente. Si $x = (x_1, \ldots, x_n, \ldots)$, $y = (y_1, \ldots, y_n, \ldots) \in E$, donde $x_i, y_i \in E_i$ ($i = 1, \ldots, n, \ldots$),

$\lambda \in K$, definamos $x + y = (x_1 + y_1, \ldots, x_n + y_n, \ldots) \in E$, $\lambda x = (\lambda x_1, \ldots, \lambda x_n, \ldots) \in E$. Por último, de un modo general, probar que el producto cartesiano $E = \prod_{i \in I} E_i$ de los espacios vectoriales E_i $(i \in I)$ sobre el cuerpo K, donde I es un conjunto no vacío, es un espacio vectorial sobre K del siguiente modo natural. Si $x = (x_i)_{i \in I}$, $y = (y_i)_{i \in I} \in E$, donde $x_i, y_i \in E_i (i \in I)$, $\lambda \in K$, definamos $x + y = (x_i + y_i)_{i \in I} \in E$, $\lambda x = (\lambda x_i)_{i \in I} \in E$. En particular, si E fuese un espacio vectorial e I un conjunto no vacío dado, la potencia cartesiana E^I es un espacio vectorial sobre K.

§2. SUBESPACIOS VECTORIALES

El concepto de subespacio vectorial desempeña en la teoría de los espacios vectoriales un papel semejante al de la noción de subanillo en la teoría de los anillos, o de subgrupo en la teoría de grupos, o de subconjunto en la teoría de los conjuntos. Lo más usual es que cuando se consideran dos espacios vectoriales E y F sobre el mismo cuerpo K, tales que F sea un subconjunto de E, las operaciones de adición y de multiplicación en E actúan del mismo modo sobre los elementos de F: esto significa que, si $x, y \in F$ (luego, $x, y \in E$), entonces la suma $x + y$ obtenida cuando se considera x e y como elementos de F, coincide con la suma $x + y$ de x e y considerados como elementos de E; además de lo cual, si $\lambda \in K$ y $x \in F$ (luego $\lambda \in K$, $x \in E$), el producto λx tiene el mismo valor cuando se piensa en x como elemento de F o como elemento de E. Tal concepto será ilustrado en los ejemplos siguientes.

Un subconjunto F de un espacio vectorial E sobre un cuerpo K es un *subespacio vectorial* de E sobre K cuando F es un espacio vectorial sobre K con respecto a las operaciones de adición y de multiplicación en E restringidas a F. Esto significa que F es un espacio vectorial sobre K con respecto a las dos correspondencias $(x, y) \mapsto x + y$ y $(\lambda, x) \mapsto \lambda x$, donde $x, y \in F$, $\lambda \in K$, y se calculan $x + y$ y λx como en E. Por lo general, se dice simplemente que F es un subespacio vectorial de E, sin aludir al cuerpo K, cuando se trata del mismo cuerpo de escalares para E y F.

Proposición 1. Para que el subconjunto F del espacio vectorial E sobre el cuerpo K sea un subespacio vectorial es necesario y suficiente que:

1. F sea un subgrupo aditivo de E.
2. $\lambda \in K$ y $x \in F$ impliquen $\lambda x \in F$.

La demostración es análoga a la de la proposición, 1, §3, capítulo 3 y, por tanto, será omitida.

Notemos que en general se verifica la condición 1 al constatar, teniendo en cuenta la condición 2, que:

1'. $0 \in F$, 1''. $x, y \in F$ implican $x + y \in F$.

En efecto, la condición 1 implica las condiciones 1' y 1'', por la proposición 1, §3, capítulo 2. Recíprocamente, teniendo en cuenta la condición 2 (la cual acarrea $-x = (-1)x \in F$, si $x \in F$) vemos que las condiciones 1' y 1'' implican la condición 1 por la misma proposición 1, §3, capítulo 2.

Ejemplo 1. Consideremos el espacio vectorial $F = F(\mathbf{R};\mathbf{R})$ de las funciones reales de variable real, esto es las funciones definidas en $\mathbf{R}$ con valores en $\mathbf{R}$ (véase el ejemplo 3, § 1, capítulo 4 y el comentario al final). Si consideramos el conjunto $C = C(\mathbf{R};\mathbf{R})$ de las funciones reales continuas de variable real, entonces C es un subespacio vectorial de F. Si consideramos el conjunto $P = P(\mathbf{R};\mathbf{R})$ de los polinomios reales de variable real, entonces P es un subespacio vectorial de C. Si consideramos el conjunto $P_m = P_m(\mathbf{R};\mathbf{R})$ de los polinomios reales de variable real y de grado no mayor que $m = 0, 1, 2, \ldots,$ entonces P_m es un subespacio vectorial de P. Observemos que el conjunto de los polinomios reales de variable real y de grado exactamente igual a m es un subconjunto de P_m, mas no un subespacio vectorial de P_m. Las mismas consideraciones de este ejemplo se repiten sustituyendo $\mathbf{R}$ por $\mathbf{C}$.

Es obvio que un espacio vectorial E es un subespacio de sí mismo y que el conjunto reducido al 0 de E es un subespacio vectorial de E. Así, pues, E es el más grande subespacio vectorial de sí mismo y 0 es el menor de los subespacios vectoriales de E.

En el caso de un grupo aditivo, se ha definido ya la notación $X + Y$, donde X e Y son dos partes cualesquiera del grupo (pág. 59). Si E es un espacio vectorial sobre un cuerpo K, definamos igualmente

$$\Lambda X = \{\lambda x;\ \lambda \in \Lambda,\ x \in X\},$$

donde $\Lambda \subset K$ y $X \subset E$, o sea ΛX es el conjunto de los elementos de E de la forma λx, donde λ y x varían en Λ y X, respectivamente. Podemos, entonces, reformular la proposición 1 precedente de la siguiente manera:

Proposición 2. Para que el subconjunto F del espacio vectorial E sea un subespacio vectorial, es necesario y suficiente que:

$$1.\ 0 \in F, \quad 2.\ F + F \subset F, \quad 3.\ KF \subset F.$$

Ejercicios

1) Sea E un espacio vectorial sobre el cuerpo K. Si $F_1, \ldots, F_n$ son subespacios vectoriales de E, probar que la intersección $F = F_1 \cap \ldots \cap F_n$ es un subespacio vectorial de E. Si $F_1, \ldots, F_n, \ldots$ son subespacios vectoriales de E, probar que la intersección $F = F_1 \cap \ldots \cap F_n \cap \ldots$ es un subespacio vectorial de E. Finalmente, de modo más general, si F_i ($i \in I$) son subespacios vectoriales de E, donde I es un conjunto no vacío, probar que la intersección $F = \bigcap_{i \in I} F_i$ es un subespacio vectorial de E.

2) Sea E un espacio vectorial sobre el cuerpo K. Si $F_1 \subset \ldots \subset F_n \subset \ldots$ son subespacios vectoriales de E, probar que la unión $F = F_1 \cup \ldots \cup F_n \cup \ldots$ es un subespacio vectorial de E. De un modo más general, probar que la unión $F = \bigcup_{i \in I} F_i$ es un subespacio vectorial de E, si F_i ($i \in I$) son subespacios vectoriales de E, donde I es un conjunto no vacío que cumple la siguiente condición: si i_1 e $i_2 \in I$, existe un $i_3 \in I$, tal que $F_{i_1} \subset F_{i_3}$, $F_{i_2} \subset F_{i_3}$.

3) Sea E un espacio vectorial sobre un cuerpo K. Si $F_1, \ldots, F_n$ son subespacios vectoriales de E, probar que la suma $F = F_1 + \ldots + F_n$ es

un subespacio vectorial de E. Si $F_1, \ldots, F_n, \ldots$ son subespacios vectoriales de E, probar que la suma $F = \sum_{n=1}^{\infty} F_n = \bigcup_{n=1}^{\infty} (F_1 + \ldots + F_n)$ es un subespacio vectorial de E. Finalmente, de un modo más general, probar que la suma $F = \sum_{i \in I} F_n = \bigcup (F_{i_1} + \ldots + F_{i_n})$ es un subespacio vectorial de E, si F_i $(i \in I)$ son subespacios vectoriales de E, donde I es un conjunto no vacío y la unión indicada se toma para todos los valores de $n = 1, 2, \ldots, i_1, \ldots, i_n \in I$.

4) Si F, F_1, F_2 son subespacios vectoriales de un espacio vectorial E, entonces $F \subset F_1 \cup F_2$ si, y sólo si, $F \subset F_1$ o $F \subset F_2$. Concluir que la unión $F_1 \cup F_2$ de dos subespacios vectoriales F_1 y F_2 de un espacio vectorial E es un subespacio vectorial de E si, y sólo si, $F_1 \subset F_2$ o $F_2 \subset F_1$. Generalizar estos dos enunciados a los casos de subespacios vectoriales F, $F_1, \ldots, F_n$, así como de subespacios vectoriales $F_1, \ldots, F_n$, de un espacio vectorial E.

5) Si F, G_1, G_2 son subespacios vectoriales de un espacio vectorial E, mostrar que $F \cap G_1 = F \cap G_2$, $F + G_1 = F + G_2$ no implican, en general, $G_1 = G_2$, pero que esto ocurre si $G_1 \subset G_2$.

6) Si F, G, H son subespacios de un espacio vectorial E, mostrar que $F \cap (G + H) = (F \cap G) + (F \cap H)$ puede ser falso; pero que $F \cap (G + F \cap \cap H) = (F \cap G) + (F \cap H)$ es verdadero.

§ 3. APLICACIONES LINEALES E ISOMORFISMOS

Dados dos espacios vectoriales E y F sobre el mismo cuerpo K, se llama *aplicación lineal u homomorfismo* de E en F a toda función $f: E \to F$ tal que

$$f(x + y) = f(x) + f(y), \quad f(\lambda x) = \lambda f(x) \quad (x, y \in E, \lambda \in K),$$

o sea a toda aplicación de E en F que respete las operaciones del espacio vectorial. Una aplicación lineal de E en F es, en particular, un homomorfismo de grupo aditivo según la primera de las condiciones precedentes, $f(x + y) = f(x) + f(y)$. Todos los hechos válidos para los homomorfismos de grupos aditivos, como

$$f(0) = 0, \quad f(-x) = -f(x),$$

son, por lo tanto, automáticamente válidos para las aplicaciones lineales de espacios vectoriales. Notemos solamente, a título de curiosidad, que estas dos igualdades pueden probarse vectorialmente, haciendo $\lambda = 0$ y $\lambda = -1$ en $f(\lambda x) = \lambda f(x)$.

Un *isomorfismo* del espacio vectorial E en el espacio vectorial F es, por definición, cualquier aplicación lineal biunívoca de E en F. Un isomorfismo *entre* E y F es, conforme a una convención anterior (correspondencia biunívoca entre conjuntos, pág. 16; isomorfismo entre grupos, pág. 67; isomorfismo entre anillos, pág. 95) un isomorfismo de E sobre F. Dos espacios vectoriales se dicen *isomorfos* y se consideran idénticos desde el punto de vista abstracto --esto es, cuando lo que tiene importancia no es la naturaleza de los elementos que constituyen un espacio vectorial, sino el modo por el cual se combinan algebrai-

camente-- cuando existe al menos un isomorfismo entre estos espacios vectoriales.

De un modo general, diremos que un espacio vectorial F es imagen lineal del espacio vectorial E cuando existe por lo menos una aplicación lineal de E sobre F.

Como en los casos de los grupos aditivos y de los anillos, la función constante $0:E \to F$, que a cada elemento de E asocia siempre el cero de F, es una aplicación lineal, llamada aplicación lineal cero de E en F (véase la pág. 13, en lo que atañe a las funciones constantes).

A toda aplicación lineal $f:E \to E$ de un espacio vectorial en sí mismo se le da el nombre de *endomorfismo*. Se llama *automorfismo* de un espacio vectorial E a cualquier isomorfismo de E sobre sí mismo. La transformación identidad $I:E \to E$ es un ejemplo evidente de automorfismo. Notemos que si fijamos $\mu \in K$, la aplicación $x \in E \mapsto \mu x \in E$ es lineal cuando K es conmutativo, lo que significa que $\lambda\mu = \mu\lambda$ para todo $\lambda \in K$; esta aplicación es un automorfismo de E cuando $\mu \neq 0$.

Una función lineal $f:E \to K$ se llama *forma lineal;* es el caso de una aplicación lineal $f : E \to F$ cuando el espacio vectorial F de los valores es el propio cuerpo de los escalares, $F = K$.

Ejemplo 1. Consideremos el espacio vectorial $P = P(\mathbf{R}, \mathbf{R})$ de los polinomios reales de variable real (véase el ejemplo 1, § 2, capítulo 4). Si fijamos un polinomio $p \in P$, entonces la operación de multiplicación $f \in P \mapsto pf \in P$ es lineal. Supongamos que p sea de grado g. La aplicación de multiplicación $f \in P_m \mapsto pf \in P_{g+m}$ es lineal (véase la notación en el ejemplo citado arriba). La derivación $f \in P \mapsto f' \in P$ es una aplicación lineal y lo mismo vale para la derivación $f \in P_m \mapsto f' \in P_{m-1}$, si $m \geq 1$. Si se considera el espacio $C = C[a, b]$ de las funciones reales continuas en $[a, b] \subset \mathbf{R}$, entonces la función $f \in C \mapsto \int_a^b f(x)\, dx \in \mathbf{R}$ es una forma lineal.

113

Proposición 1. Si $f:E \to F$ y $g:F \to G$ son aplicaciones lineales entre los espacios vectoriales E, F y G, entonces $gf:E \to G$ es una aplicación lineal.

La demostración es análoga a la de la proposición 1, § 4, capítulo 2 y se omite.

Proposición 2. Si $f:E \to F$ es un isomorfismo entre los espacios vectoriales E y F, entonces $f^{-1}:F \to E$ es un isomorfismo entre F y E.

La demostración es similar a la de la proposición 2, § 4, capítulo 2 y se omite.

Ejercicios

1) Sea $f:E \to F$ una aplicación lineal entre los espacios vectoriales E y F. Si X es un subespacio vectorial de E, entonces la imagen directa

$f(X)$ es un subespacio vectorial de F. En particular, $f(E)$ es un subespacio vectorial de F, o sea el conjunto de los puntos de F de la forma $y = f(x)$, para todo $x \in E$, es un subespacio vectorial de E. Si Y fuese un subespacio vectorial de F, entonces la imagen inversa $f^{-1}(Y)$ será un subespacio vectorial de E. En particular, $f^{-1}(0)$ es un subespacio vectorial de E, o sea, el conjunto de los puntos x de E que satisfacen la ecuación $f(x) = 0$ es un subespacio vectorial de E.

2) Dada la aplicación $f: E \to F$ entre los espacios vectoriales E y F, consideremos el gráfico G de f (§ 9, capítulo 1, pág. 35), que es un subconjunto del espacio vectorial $E \times F$ (ejercicio 7, § 1, capítulo 4). Probar que f es una àplicación lineal si, y sólo si, G es un subespacio vectorial de $E \times F$.

3) Si $E = E_1 \times \ldots \times E_n$ fuese el producto cartesiano de los espacios vectoriales $E_1, \ldots, E_n$, probar que cada proyección $\pi_i: E \to E_i$ ($i = 1, \ldots, n$) (§ 9, capítulo 1, pág. 35) es una aplicación lineal de E sobre E_i. Análogamente, si $E = E_1 \times \ldots \times E_n \times \ldots$ fuese el producto cartesiano de los espacios vectoriales $E_1, \ldots, E_n, \ldots$, probar que cada proyección $\pi_i: E \to E_i$ ($i = 1, \ldots, n, \ldots$) es una aplicación lineal de E sobre E_i.

4) Si E y F son espacios vectoriales sobre el cuerpo conmutativo o campo K, designemos por $\mathcal{L}(E; F)$ al conjunto de las aplicaciones lineales de E en F. Probar que $\mathcal{L}(E; F)$ es un espacio vectorial con respecto a las siguientes operaciones de adición y multiplicación: si $f, g \in \mathcal{L}(E; F)$, $\lambda \in K$, entonces $f + g$, $\lambda f \in \mathcal{L}(E; F)$ quedan definidas por $(f + g)(x) = f(x) + g(x)$ y $(\lambda f)(x) = \lambda f(x)$, para todo $x \in E$. En particular, si E es un espacio vectorial sobre un cuerpo conmutativo K, el conjunto E^* de las formas lineales sobre E es un espacio vectorial con respecto a la adición y multiplicación precedentes (se toma $F = K$); E^* se denomina *espacio vectorial dual* de E.

5) Si E_1 y E_2 son subespacios vectoriales del espacio vectorial E, se dice que E es la suma directa de E_1 y E_2, y se escribe $E = E_1 \oplus E_2$, cuando todo elemento $x \in E$ puede ser escrito de manera única como $x = x_1 + x_2$, donde $x_1 \in E_1$, $x_2 \in E_2$. Probar que E es la suma directa de E_1 y E_2 si, y sólo si, $E_1 + E_2 = E$, $E_1 \cap E_2 = \{0\}$. De un modo más general, si $E_1, \ldots, E_n$ ($n \geq 2$) son subespacios vectoriales del espacio vectorial E, se dice que E es la suma directa de $E_1, \ldots, E_n$ y se escribe $E = E_1 \oplus \ldots \oplus E_n$, cuando todo elemento $x \in E$ puede ser escrito de forma única como la suma $x = x_1 + \ldots + x_n$, donde $x_i \in E_i$ ($i = 1, \ldots, n$). Probar que E es la suma directa de $E_1, \ldots, E_n$ si, y sólo si (suponiendo, por razones de claridad, $n \geq 3$), $E_1 + \ldots + E_n = E$, $E_1 \cap (E_2 + \ldots + E_n) = \{0\}$, $E_2 \cap (E_3 + \ldots + E_n) = \{0\}, \ldots, E_{n-1} \cap E_n = \{0\}$ (notar que estas condiciones dependen del orden en que los espacios $E_1, \ldots, E_n$ sean enumerados, de modo que se obtienen condiciones distintas, pero equivalentes, enumerándolos de manera diferente). De esto resulta que E es la suma directa de $E_1, \ldots, E_n$ si, y sólo si, $E_1 + \ldots + E_n = E$, $E_i \cap (E_1 + \ldots + \hat{E}_i + \ldots + E_n) = \{0\}$ ($i = 1, \ldots, n$), donde $n \geq 2$ y el término $\hat{E}_i$ de orden i de la suma debe ser omitido (adviértase que estas condiciones no dependen del orden en que se enumeren los espacios $E_1, \ldots, E_n$, en el sentido de que ellos desempeñan el mismo papel en ellas). La noción de suma directa $E = \oplus_{i \in I} E_i$ puede extenderse a una familia de subespacios vectoriales E_i ($i \in I$) de un espacio vectorial E, donde I es un conjunto no vacío.

6) Consideremos el espacio vectorial K^n sobre el cuerpo K, donde $n = 1, 2, \ldots$ (ejercicio 7, § 1, capítulo 4). Una vez fijados $a_1, \ldots, a_n \in$ $\in K$, definamos $f: K^n \to K$ por $f(x) = x_1a_1 + \ldots + x_na_n \in K$ (que se acostumbra escribir $a_1x_1 + \ldots + a_nx_n$ cuando K es conmutativo), para todo $x =$ $= (x_1, \ldots, x_n) \in K^n$, $x_i \in K (i = 1, \ldots, n)$. Probar que f es una forma lineal, esto es $f \in (K^n)^*$ (ejercicio 4, § 3, capítulo 4), y que, recíprocamente, toda forma lineal f sobre K^n se representa de la manera indicada, de modo único, o sea para $a_i \in K$ $(i = 1, \ldots, n)$ únicos. Mostrar que esta correspondencia biunívoca $f \in (K^n)^* \mapsto (a_1, \ldots, a_n) \in K^n$, del primer espacio vectorial sobre el segundo, es aditiva y que ella es lineal si K es conmutativo.

7) Consideremos los espacios vectoriales K^n y K^m sobre el cuerpo K, $n, m = 1, 2, \ldots$ (ejercicio 7, § 1, capítulo 4). Una vez fijada la matriz

$$\begin{pmatrix} a_{11} & & a_{1n} \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$$

de m filas y n columnas, donde $a_{ij} \in K$ $(i = 1, \ldots, m, j = 1, \ldots, n)$, definamos $f: K^n \to K^m$ por $f(x) = (y_1, \ldots, y_m) \in K^m$, $y_i = \sum_{j=1}^{n} x_j a_{ij}$ $(i = 1, \ldots, m)$, (donde se acostumbra escribir $a_{ij}x_j$ si K es conmutativo), para todo $x = (x_1, \ldots, x_n) \in K^n$, $x_i \in K (i = 1, \ldots, n)$. Probar que f es una aplicación lineal y que, recíprocamente, toda aplicación lineal f de K^n en K^m se representa de la manera indicada, de modo único, o sea para $a_{ij} \in$ $\in K$ $(i = 1, \ldots, m, j = 1, \ldots, n)$ únicos. Se puede completar este ejercicio con una aserción análoga a la del final del ejercicio precedente.

§ 4. ESPACIOS VECTORIALES DE DIMENSION FINITA

Históricamente, los primeros espacios vectoriales considerados fueron $\mathbf{R}^n$ y $\mathbf{C}^n$ (casos particulares importantes de K^n, donde K es un cuerpo) de dimensión finita n, donde n es el número entero que figura explícitamente en la definición de los mismos $(n = 1, 2, \ldots)$. Tales espacios vectoriales tienen una base natural, formada por los vectores $e_i =$ $= (0, \ldots, 1, \ldots, 0)$ $(i = 1, \ldots, n)$, cuyas coordenadas son todas el cero del cuerpo, salvo la i-ésima coordenada que es la unidad del cuerpo. De esta base se obtiene un sistema natural de coordenadas del espacio vectorial. Posteriormente, la atención se dirigió a otros espacios vectoriales que llamamos de dimensión infinita (por no ser reducibles a K^n con n un entero natural), formados sobre todo por funciones, o sea los llamados espacios funcionales. La definición axiomática (abstracta o intrínseca) de un espacio vectorial sobre un cuerpo (§ 1, capítulo 4) es relativamente reciente. Ella nos permite, de modo más satisfactorio, distinguir entre los espacios vectoriales de dimensión finita y los de dimensión infinita; para estos últimos también es posible hablar de la dimensión de cada uno de ellos, la cual es un número cardinal infinito, lo que no abordaremos, sino que se deja para un estudio algebraico posterior más profundo. A pesar de la importancia indiscutible de los espacios vectoriales de dimensión infinita, tanto en matemática como en sus aplicaciones, son los espacios vectoriales de dimensión finita los que se presentan con el derecho de primacía en el estudio inicial del

álgebra lineal, por presentar una enorme riqueza conceptual, técnica y proposicional.

Sea, pues, E un espacio vectorial sobre un cuerpo K. Diremos que $x \in E$ es una *combinación lineal* de $x_1, \ldots, x_n$, cuando existen $\lambda_1, \ldots, \lambda_n \in K$ tales que $x = \lambda_1 x_1 + \ldots + \lambda_n x_n$. Notemos que tales coeficientes escalares pueden no ser únicos.

Proposición 1. Si $x_1, \ldots, x_n$ son vectores dados en un espacio vectorial E sobre un cuerpo K, las siguientes condiciones son equivalentes:

1. Todo $x \in E$, que sea una combinación lineal de $x_1, \ldots, x_n$, tiene una expresión única $x = \lambda_1 x_1 + \ldots + \lambda_n x_n$, donde $\lambda_1, \ldots, \lambda_n \in K$.

2. Si $\lambda_1 x_1 + \ldots + \lambda_n x_n = 0$, donde $\lambda_1, \ldots, \lambda_n \in K$, entonces $\lambda_1 = \ldots = \lambda_n = 0$.

Demostración. Mostremos que la condición 2 implica la condición 1. En efecto, supongamos que $x = \lambda_1 x_1 + \ldots + \lambda_n x_n = \mu_1 x_1, + \ldots + \mu_n x_n$, donde $\lambda_i, \mu_i \in K$ $(i = 1, \ldots, n)$. De ahí resulta que $(\mu_1 - \lambda_1)x_1 + \ldots + (\mu_n - \lambda_n)x_n = 0$. Por la condición 2 obtenemos, entonces, $\lambda_1 = \mu_1, \ldots, \lambda_n = \mu_n$. Recíprocamente, probemos que la condición 1 implica la condición 2. Ahora, si $\lambda_1, \ldots, \lambda_n \in K$ y $0 = \lambda_1 x_1 + \ldots + \lambda_n x_n$, tenemos así una expresión del cero de E como combinación lineal de $x_1, \ldots, x_n$. Por otra parte, el cero de E siempre tiene la siguiente expresión $0 = 0\, x_1 + \ldots + 0\, x_n$ como combinación lineal de $x_1, \ldots, x_n$. Por la condición 1 obtenemos $\lambda_1 = \ldots = \lambda_n = 0$. QED

Diremos que los vectores $x_1, \ldots, x_n \in E$ como en la proposición 1 precedente son *linealmente independientes* cuando satisfacen las condiciones equivalentes 1 y 2 de su enunciado; en el caso contrario, se denominan *linealmente dependientes*, o sea cuando existen $\lambda_1, \ldots, \lambda_n \in K$, entre los cuales al menos uno es diferente del cero de K, tales que $\lambda_1 x_1 + \ldots + \lambda_n x_n = 0$. Notemos que, en la independencia o dependencia lineal de $x_1, \ldots, x_n$, el orden según se enumeran es irrelevante; o sea si σ es una permutación del conjunto $\{1, \ldots, n\}$, entonces $x_1, \ldots, x_n$ serán linealmente independientes, o dependientes, al mismo tiempo que $x_{\sigma(1)}, \ldots, x_{\sigma(n)}$. Tal observación se aplica a otros aspectos y no será repetida, salvo cuando sea esencial.

Proposición 2. Sea E un espacio vectorial sobre un cuerpo K. Entonces:

1. $x \in E$ es linealmente independiente, o linealmente dependiente, conforme $x \neq 0$ o $x = 0$.

2. Si $x \in E$, $\lambda, \mu \in K$, entonces λx y μx son linealmente dependientes.

3. Si $x_1, \ldots, x_n \in E$ fueran linealmente independientes, entonces $x_i \neq 0$ para todo i, como así también $x_i \neq x_j$ cualesquiera que sean $i \neq j$ (dado que $n \geq 2$); luego, si algún $x_i = 0$, o si $x_i = x_j$ para algún $i \neq j$ (caso $n \geq 2$), entonces $x_1, \ldots, x_n$ son linealmente dependientes.

4. Si $x_1, \ldots, x_n \in E$ fueran linealmente independientes y $1 \leq i_1 < \ldots < i_k \leq n$ fueran k enteros (donde $k = 1, \ldots, n$ y $n \geq 2$), entonces $x_{i_1}, \ldots, x_{i_n}$ son linealmente independientes; luego, si $x_{i_1}, \ldots, x_{i_n}$ fueran linealmente dependientes, entonces $x_1, \ldots, x_n$ serían linealmente dependientes.

La demostración es trivial y se omite.

Dados $x_1, \ldots, x_n$ en el espacio vectorial E sobre el cuerpo K, el conjunto F de todas las combinaciones lineales $\lambda_1 x_1 + \ldots + \lambda_n x_n$ ($\lambda_1, \ldots, \lambda_n \in K$) de $x_1, \ldots, x_n$ es evidentemente un subespacio vectorial de E, denominado el *subespacio vectorial generado* por $x_1, \ldots, x_n$, vectores éstos que llamamos *generadores* de F. Como todo subespacio vectorial de E que contenga $x_1, \ldots, x_n$ debe contener forzosamente a F, expresamos este hecho diciendo que F es el menor subespacio vectorial de E que contiene $x_1, \ldots, x_n$. Podemos decir que F es la intersección de todos los subespacios vectoriales de E que contienen $x_1, \ldots, x_n$. Por ejemplo, el subespacio vectorial de E generado por $x \in E$ es $Kx = \{\lambda x;\ \lambda \in K\}$; él se reduce o no al origen de E según sea $x = 0$ o $x \neq 0$.

Diremos que $x_1, \ldots, x_n \in E$ forman una *base* del espacio vectorial E sobre el cuerpo K cuando todo $x \in E$ puede expresarse como combinación lineal $x = \lambda_1 x_1 + \ldots + \lambda_n x_n$ ($\lambda_1, \ldots, \lambda_n \in K$) de $x_1, \ldots, x_n$, de modo único; o sea, cuando E es generado por $x_1, \ldots, x_n$, los cuales son linealmente independientes.

Ejemplo 1. Sea K un cuerpo. El espacio vectorial K^n ($n = 1, 2, \ldots$) tiene la siguiente base, llamada base natural de K^n, formada por los vectores $e_i = (0, \ldots, 1, \ldots, 0)$ ($i = 1, \ldots, n$), cuyas coordenadas son todas el cero del cuerpo, salvo la i-ésima coordenada que es la unidad del cuerpo. En efecto, si $x = (x_1, \ldots, x_n) \in K^n$, donde $x_i \in K$ ($i = 1, \ldots, n$), entonces $x = x_1 e_1 + \ldots + x_n e_n$ es una combinación lineal de $e_1, \ldots, e_n$ igual a x. Además, si $x = \lambda_1 e_1 + \ldots + \lambda_n e_n$ ($\lambda_1, \ldots, \lambda_n \in K$), entonces $x = (\lambda_1, \ldots, \lambda_n)$, de donde $\lambda_1 = x_1, \ldots, \lambda_n = x_n$, o sea x se expresa como combinación lineal de $e_1, \ldots, e_n$ de modo único.

Cabe indicar que, en vez de *base* deberíamos haber dicho *base finita ordenada* (puesto que estamos considerando una base formada por un número finito de vectores enumerados en algún orden), lo que no causará confusión en este texto. En rigor, podemos hablar de una base finita, así como de una base infinita, sin un orden prefijado, en lo que no nos detendremos.

Proposición 3. Un espacio vectorial E sobre un cuerpo K es isomorfo a un espacio vectorial K^n, si, y sólo si, E tiene una base con n elementos. Si $f: K^n \to E$ es una aplicación lineal, entonces f será un isomorfismo entre K^n y E si, y sólo si (indicando con e_i ($i = 1, \ldots, n$) a la base natural de K^n) $f(e_1), \ldots, f(e_n)$ es una base de E. Recíprocamente, si $x_1, \ldots, x_n \in E$ es una base de E, existe un isomorfismo f, y sólo uno, entre K^n y E tal que $f(e_i) = x_i$ ($i = 1, \ldots, n$).

Esta proposición resulta inmediatamente de la proposición siguiente, al usar K^n en vez de E, así como E en vez de F, además de otros detalles de notación.

Proposición 4. Sean E y F espacios vectoriales sobre el cuerpo K. Si E tiene una base $x_1, \ldots, x_n$, entonces F es isomorfo a E si, y sólo si, F tiene una base con n elementos. Si $f: E \to F$ es una aplicación lineal, entonces f será un isomorfismo entre E y F si, y sólo si, $f(x_1), \ldots, f(x_n)$ es una base de F. Recíprocamente, si $y_1, \ldots, y_n \in F$ es una base de F, existe un isomorfismo f, y sólo uno, entre E y F tal que $f(x_i) = y_i$ ($i = 1, \ldots, n$).

La demostración es bastante sencilla y se omite.

Nuestro objetivo inmediato es definir la dimensión de un espacio vectorial, cuando ella es finita; para esto necesitamos de la siguiente proposición, interesante por sí misma.

Proposición 5. Sea E un espacio vectorial sobre un cuerpo K. Si $x_1, \ldots, x_m \in E$ son generadores de E e $y_1, \ldots, y_n \in E$ son linealmente independientes en E, entonces $m \geq n$.

Demostración. Comencemos con una observación. Dado que E es generado por $x_1, \ldots, x_n \in E$, si $y \in E$, $y \neq 0$, entonces E es también generado por los m vectores $y, x_1, \ldots, \hat{x}_i, \ldots, x_m$, donde $\hat{x}_i$ indica que x_i se omite en la emuneración e $i = 1, \ldots, n$ se escoge convenientemente. En efecto, $y = \lambda_1 x_1 + \ldots + \lambda_m x_m$ para ciertos $\lambda_1, \ldots, \lambda_m \in K$. Como $y \neq 0$, entonces para algún i, $\lambda_i \neq 0$, de donde resulta que x_i es una combinación lineal de y y de los demás x_j ($j = 1, \ldots, m, j \neq i$). Luego, E es generado por y y por los demás x_j($j = 1, \ldots, m, j \neq i$).

Demostremos ahora la proposición por inducción sobre m. Ella se cumple para $m = 1$; en efecto, como E es generado por x_1 y dado que $y_1 \neq 0$, la observación precedente muestra que E es generado por $y_1, \hat{x}_1$, o sea simplemente por y_1, lo que exige $n = 1$, pues $y_1, \ldots, y_n$ son linealmente independientes. Supongamos ahora $m \geq 2$ y que la proposición se cumple para $m - 1$. Si $n < 2$, entonces $n < 2 \leq m$, y no hay nada que probar. Supongamos $n \geq 2$; como E es generado por $x_1, \ldots, x_m$ y dado que $y_1 \neq 0$, la observación precedente muestra, después de una renumeración de $x_1, \ldots, x_m$, que E es generado por $y_1, x_1, \ldots, x_{m-1}, \hat{x}_m$, o sea por $y_1, x_1, \ldots, x_{m-1}$ lo que exige, por la hipótesis de inducción, que $n - 1 \leq m - 1$, o $n \leq m$, dado que $y_2, \ldots, y_n$ son linealmente independientes. QED

Proposición 6. Si $x_1, \ldots, x_m \in E$ e $y_1, \ldots, y_n \in E$ son bases del espacio vectorial E, entonces $m = n$.

Es lo que resulta de inmediato de la proposición precedente aplicada dos veces simétricamente.

De acuerdo con la proposición 6 definimos la dimensión de un espacio vectorial E que tenga una base $x_1, \ldots, x_n$ como el entero $n = 1, 2, \ldots$, el que es independiente de la base escogida en E. Conviene notar que tal definición implica que E no consiste sólo del origen, pues $n \geq 1$ y cada $x_i \neq 0$. Completemos la definición con la convención de que la dimensión de un espacio que consiste sólo del origen, es 0. Indicaremos con dim E la dimensión de E.

Pasemos a enumerar algunas propiedades útiles de la dimensión.

Proposición 7. Si $E \neq 0$ es un espacio vectorial y $x_1, \ldots, x_n \in E$ son generadores de E, entonces E tiene dimensión $m \leq n$. Más precisamente, podemos escoger enteros $1 \leq i_1 \leq \ldots \leq i_m \leq n$ de modo que $x_{i_1}, \ldots, x_{i_m}$ sea una base de E.

Demostración. Si $x_1, \ldots, x_n$ son linealmente independientes forman una base y la proposición es evidente. Supongamos $x_1, \ldots, x_n$ linealmente dependientes; entonces $n \geq 2$ (pues $E \neq 0$) y existe x_i que es una combinación lineal de los x_j ($j \neq i$). De ahí resulta que $x_1, \ldots, \hat{x}_i, \ldots, x_n$ generan E, donde $\hat{x}_i$ indica que x_i ha sido omitido. Si $x_1, \ldots, \hat{x}_i, \ldots, x_n$ son linealmente independientes y forman una base de E, la proposición queda probada. Si suponemos $x_1, \ldots, \hat{x}_i, \ldots, x_n$ linealmente dependientes, entonces $n \geq 3$ (pues $E \neq 0$), y razonamos como en el caso anterior y así sucesivamente. Este proceso de omisiones sucesivas tiene que parar a lo más en $n - 1$ etapas, dado que $n \geq 2$ (pues $E \neq 0$), cuando la proposición queda probada. QED

Proposición 8. Si el espacio vectorial E tiene dimensión $n \geq 1$, para que $x_1, \ldots, x_n \in E$ formen una base de E es necesario y suficiente que sea satisfecha por lo menos una de las dos condiciones siguientes: 1. $x_1, \ldots, x_n$ son generadores de E; 2. $x_1, \ldots, x_n$ son linealmente independientes.

Demostración. Cada una de las dos condiciones es necesaria (además, las dos condiciones juntas significan que $x_1, \ldots, x_n$ forman una base de E). Probemos la suficiencia de la condición 1, para lo cual basta mostrar que $x_1, \ldots, x_n$ son linealmente independientes. En caso contrario, se podría omitir un x_i, de modo que $x_1, \ldots, \hat{x}_i, \ldots, x_n$ generen E, donde $\hat{x}_i$ indica que x_i fue omitido; pero, entonces, la proposición 7 implicaría que la dimensión de E es a lo sumo $n - 1$, lo que no es cierto. Probemos la suficiencia de la condición 2, para lo cual basta mostrar que $x_1, \ldots, x_n$ son generadores de E. En caso contrario, $x_1, \ldots, x_n$ generan un subespacio vectorial propio F de E y podemos escoger $x_{n+1} \in E$, $x_{n+1} \notin F$; entonces $x_1, \ldots, x_n, x_{n+1}$ son $n + 1$ vectores linealmente independientes de E, en tanto que una base de E está formada por n generadores de E, lo que contradice la proposición 5. QED

Proposición 9. Si el espacio vectorial E tiene dimensión m y $x_1, \ldots, x_n \in E$ son linealmente independientes, entonces $m \geq n$. Cuando $m = n$, $x_1, \ldots, x_n$ forman una base de E. Si $m > n$, podemos hallar $x_{n+1}, \ldots, x_m \in E$ tales que $x_1, \ldots, x_n, x_{n+1}, \ldots, x_m$ sea una base de E.

Demostración. La hipótesis de que $x_1, \ldots, x_n \in E$ sean linealmente independientes, cuando $n \geq 1$, implica que $E \neq 0$. Como una base de E está formada por m generadores de E, de la proposición 5 se desprende que $m \geq n$. Cuando $m = n$, la condición 2 de la proposición 8 muestra que $x_1, \ldots, x_n$ forman una base de E. Supongamos $m > n$. El subespacio vectorial F_1 de E generado por $x_1, \ldots, x_n$ tiene dimensión $n < m$; luego, F_1 es propio en E. Escojamos $x_{n+1} \in E$, $x_{n+1} \notin F_1$ y sea F_2 el subespacio vectorial de E generado por $x_1, \ldots, x_n, x_{n+1}$ que tiene dimensión $n + 1 \leq m$. Cuando $m = n + 1$, la condición 2 de la proposición 8 muestra que $x_1, \ldots, x_n, x_{n+1}$ forman una base de E. Suponiendo $m > n + 1$ y continuando de esta forma, la condición 2 de la proposición 8 muestra que pasaremos por $m - n$ etapas, antes de completar esta demostración. QED

Proposición 10. Si F es un subespacio vectorial del espacio vectorial E de dimensión finita m, entonces F tiene dimensión $n \leq m$.

Demostración. Para evitar malentendidos, notemos que a partir de la hipótesis de que $E \neq 0$ tenga una base no es posible concluir que $F \neq 0$ tiene una base. Si $F = 0$ la proposición es evidente. Supongamos $F \neq 0$, luego $E \neq 0$. Escojamos $x_1 \in F$, $x_1 \neq 0$. Sea F_1 el subespacio vectorial de F generado por x_1, que es de dimensión 1. Si $F_1 = F$, entonces x_1 es una base de F, que tiene dimensión 1; la proposición 5 muestra que $1 \leq$ $\leq m$. Si F_1 es un subespacio vectorial propio de F, escojamos $x_2 \in F$, $x_2 \notin F_1$. Sea F_2 el subespacio vectorial de F generado por x_1, x_2, que es de dimensión 2. Si $F_2 = F$, entonces x_1, x_2 es una base de F, que tiene dimensión 2; la proposición 5 muestra que $2 \leq m$. Si F_2 es un subespacio vectorial propio de F continuamos de forma similar, y la proposición 5 muestra que pasaremos por $n \leq m$ etapas, construyendo una base $x_1, \ldots, x_n$ de F. QED

Ejercicios

1) Si I es un conjunto no vacío y K un cuerpo, entonces el espacio vectorial K^I de todas las funciones de I en K es de dimensión finita n si, y sólo si, I tiene n elementos.

2) Calcular la dimensión del espacio vectorial $P_m(\mathbf{R}^n; \mathbf{R})$ de los polinomios reales de n variables reales y de grado a lo sumo igual a m. Análogamente, para $\mathbf{C}$ en vez de $\mathbf{R}$, y $\mathbf{Z}/p$ en vez de $\mathbf{R}$, donde $p \geq 2$ es un entero primo.

3) Si $x_1, \ldots, x_n$ son linealmente dependientes en un espacio vectorial E, donde $x_1 \neq 0$ y $n \geq 2$, entonces algún x_i ($i = 2, \ldots, n$) es una combinación lineal de los precedentes $x_1, \ldots, x_{i-1}$.

4) Los vectores $x_1, \ldots, x_n$ ($n \geq 2$) de un espacio vectorial E son linealmente dependientes si, y sólo si, omitiéndose uno de ellos, los $n - 1$ restantes son aún generadores del mismo subespacio vectorial de E, que es generado por los n vectores dados.

5) Si E es un espacio vectorial de dimensión finita y F es un subespacio vectorial de E tal que E y F tienen la misma dimensión, entonces $E = F$.

6) En un espacio vectorial E de dimensión $n \geq 1$ no hay $m > n$ vectores linealmente independientes; luego n es el número máximo de vectores linealmente independientes en E. Por otra parte, en un espacio vectorial E de dimensión $n > 1$ no existen $m < n$ vectores que sean generadores de E; luego n es el número mínimo de generadores de E.

7) Si $f: E \to F$ es una aplicación lineal entre los espacios vectoriales E y F, donde E es de dimensión finita, entonces los subespacios vectoriales $f^{-1}(0) \subset E$ y $f(E) \subset F$ son también de dimensión finita, y se tiene que $\dim E = \dim f^{-1}(0) + \dim f(E)$. Si $f^{-1}(0)$ y $f(E)$ son de dimensión finita, también E es de dimensión finita.

8) Si $E = E_1 \times \ldots \times E_p$ es un producto cartesiano de espacios vectoriales, cada uno de dimensión n_i ($i = 1, \ldots, p$), entonces E tiene dimensión $n_1 + \ldots + n_p$ y, si E tiene dimensión finita, lo mismo ocurrirá con cada E_i ($i = 1, \ldots, p$).

9) Si $E = F_1 + \ldots + F_p$ es un espacio vectorial que es la suma de sus subespacios vectoriales $F_1, \ldots, F_p$, entonces E es de dimensión finita si, y sólo si, todos los $F_1, \ldots, F_p$ son de dimensión finita; se tiene, entonces, $\dim E \leq \dim F_1 + \ldots + \dim F_p$. Para que se cumpla la igualdad es necesario y suficiente que la suma sea directa, $E = F_1 \oplus \ldots \oplus F_p$ (ejercicio 5, § 3, capítulo 4).

10) Si F_1 y F_2 son subespacios vectoriales de dimensión finita de un espacio vectorial E, entonces $F_1 \cap F_2$ y $F_1 + F_2$ son subespacios vectoriales de E de dimensión finita; se tiene $\dim (F_1 \cap F_2) + \dim (F_1 + F_2) = \dim F_1 + \dim F_2$. Generalizar al caso de subespacios vectoriales $F_1, \ldots, F_p$ de dimensión finita de un espacio vectorial E.

11) Sean dados los enteros m, n_1, n_2, donde $m \geq n_1 \geq 0$, $m \geq n_2 \geq 0$. Hallar el mayor valor posible de $\dim (F_1 + F_2)$ y el menor valor posible de $\dim (F_1 \cap F_2)$, si tenemos un espacio vectorial E de dimensión m y subespacios vectoriales F_1 y F_2 de E, de dimensión n_1 y n_2, respectivamente. Generalizar al caso de enteros $m, n_1, \ldots, n_p$, un espacio vectorial E y sus subespacios vectoriales $F_1, \ldots, F_p$ de dimensiones $m, n_1, \ldots, n_p$, respectivamente.

12) Sea $f: E \to F$ una aplicación lineal entre los espacios vectoriales E y F de dimensiones m y n, respectivamente. Si $m > n$, entonces f no puede ser un isomorfismo de E en F. Si $m < n$, entonces f no puede ser sobre F. Si $m = n$, entonces que f sea un isomorfismo es equivalente a que f sea sobre F.

13) Si E y F son espacios vectoriales sobre un cuerpo conmutativo K, mostrar que el espacio vectorial de dimensión finita $\mathcal{L}(E;F)$ sobre K es de dimensión finita si, y sólo si, E y F son de dimensión finita, en cuyo caso $\dim \mathcal{L}(E;F) = \dim E \cdot \dim F$. Mostrar, entonces, que E^* es de dimensión finita si, y sólo si, E es de dimensión finita, en cuyo caso $\dim E^* = \dim E$ (ejercicio 4, § 3, capítulo 4).

14) Sea E un espacio vectorial real. Si $x_1, \ldots, x_n \in E$, llamamos combinación *convexa* de $x_1, \ldots, x_n$ a toda combinación lineal $x = \lambda_1 x_1 + \ldots + \lambda_n x_n$, donde $\lambda_1, \ldots, \lambda_n \in \mathbf{R}$, $\lambda_1, \ldots, \lambda_n \geq 0$, $\lambda_1 + \ldots + \lambda_n = 1$. Mostrar que toda combinación convexa de $x_1, x_2 \in E$ puede escribirse como $x = (1 - \lambda) x_1 + \lambda x_2$, donde $\lambda \in \mathbf{R}$, $0 \leq \lambda \leq 1$. Decimos que un subconjunto X de E es *convexo* cuando cualesquiera que sean $x_1, x_2 \in X$, toda combinación convexa de x_1, x_2 pertenece a X. Probar que X es convexo si, y sólo si, cualesquiera que sean $x_1, \ldots, x_n \in X$, toda combinación convexa de $x_1, \ldots, x_n$ pertenece a X. Un cono en E de *vértice* 0 es un subconjunto X de E tal que si $x \in X$, $\lambda \in \mathbf{R}$, $\lambda > 0$, entonces $\lambda x \in X$. Probar que un cono en E de vértice 0 es convexo si, y sólo si, cualesquiera que sean $x_1, x_2 \in X$ se tiene $x_1 + x_2 \in X$.

15) Sea E un espacio vectorial sobre el cuerpo K. Se llama *subespacio afín* no vacío de E a todo subconjunto X de E tal que $X = F + x$, don-

de $F \subset E$ es un subespacio vectorial y $x \in E$; entonces $x \in X$, y F queda determinado por X, pues $F = X - x$ cualquiera que sea $x \in X$, al mismo tiempo que cualquier $x \in X$ puede aparecer en la representación $X = F + x$. (Es conveniente considerar al conjunto vacío como un subespacio afín de E.) Si $x_1, \ldots, x_n \in E$, llamamos combinación *afín* de $x_1, \ldots, x_n$ a toda combinación lineal $x = \lambda_1 x_1 + \ldots + \lambda_n x_n$, donde $\lambda_1, \ldots, \lambda_n \in K$, $\lambda_1 + \ldots + \lambda_n = 1$. Mostrar que toda combinación afín de $x_1, x_2 \in E$ puede ser escrita como $x = (1 - \lambda) x_1 + \lambda x_2$, donde $\lambda \in K$. Probar que un subconjunto X de E es un subespacio afín si, y sólo si, cualesquiera que sean $x_1, x_2 \in X$, toda combinación afín de x_1, x_2 pertenece a X; o, de modo equivalente, si, y sólo si, cualesquiera que sean $x_1, \ldots, x_n \in X$, toda combinación afín de $x_1, \ldots, x_n$ pertenece a X. Mediante la representación $X = F + x$, definir la dimensión finita de X, la base afín finita de X, etc.

5

ORDEN

La noción general de orden tiene su origen tanto en matemática como en lógica. Al parecer, fue considerada por primera vez en el siglo XIX, aunque algunas de sus raíces pueden atribuirse a trabajos anteriores. El origen del importante tema del orden total, que desde el punto de vista de la matemática de hoy constituye simplemente un ejemplo particular y tan antiguo como las ideas de número, de tiempo y orientación de una recta, se pierde en el pasado. Cronológicamente, los estudios de Boole sobre el análisis de la lógica desde el punto de vista matemático, esto es, sus investigaciones sobre las leyes del pensamiento, fueron los trabajos más importantes sobre la idea de orden. Su nombre quedó así asociado a las llamadas álgebras de Boole, que durante un largo tiempo fueron el único ejemplo conocido de un sistema algebraico cuyos elementos estaban desprovistos de sentido numérico. Estas álgebras pasaron a tener interés matemático en los tiempos actuales, inclusive desde el punto de vista de sus aplicaciones a otros campos del saber. A Dedekind se debe la observación de que el concepto de conjunto ordenado aparece con gran frecuencia en matemática a tal punto de ser objeto de estudio autónomo; sin duda, el estudio de la noción general de orden es un instrumento precioso para comprender los fundamentos de varias ramas de la matemática y de sus aspectos comunes. Sin embargo, no es superfluo señalar aquí que un tal estudio no debe constituir un objetivo en sí mismo; su valor matemático se restringe a las posibilidades de sus aplicaciones. Este capítulo es una mera introducción al lenguaje del orden.

§ 1. ORDEN TOTAL

En matemática elemental se encuentran varios ejemplos de conjuntos cuyos elementos pueden compararse entre sí por una relación de orden, como lo ilustran los casos siguientes.

Ejemplo 1. Entre los números enteros naturales, o sea los elementos de **N**, existe una relación de orden usual, que se escribe $x \leq y$ para indicar que x es igual o menor que y, donde $x, y \in \mathbf{N}$. Tal relación posee las propiedades siguientes:

$$x \leq x,$$

$$x \leq y,\ y \leq z \Rightarrow x \leq z$$

$$x \leq y,\ y \leq x \Rightarrow x = y$$

$$x \leq y \text{ o } y \leq x.$$

Análogamente, para la relación de orden usual en cada uno de los conjuntos **Z**, **Q** y **R**. Observemos que en el caso del conjunto **C** no se acostumbra imponer ninguna relación de orden.

Ejemplo 2. En geometría elemental, dada una recta R, se destaca que existen dos orientaciones posibles sobre la misma, siendo cada una de ellas la opuesta de la otra. Una vez fijada una de tales orientaciones en R queda automáticamente establecida una relación de orden entre los puntos de R. En este caso, si x e y son dos puntos de R, se dice que x precede a y cuando $x = y$, o cuando $x \neq y$, y la orientación de R corresponde a un desplazamiento de un punto que se mueve a partir de x hacia y. Se escribe entonces $x \leq y$. Las cuatro propiedades indicadas en el ejemplo precedente se mantienen válidas en este ejemplo. Además, como se puede apreciar, el caso de **R** como conjunto ordenado de la forma indicada en el ejemplo anterior y el caso de la recta R como conjunto ordenado discutido en este ejemplo son esencialmente dos aspectos del mismo fenómeno. Como se ve en geometría elemental, si se fija un origen, una unidad de medida y una orientación en la recta R se obtiene un sistema de coordenadas en R y una correspondencia biunívoca entre **R** y R que respeta el orden (esto es, un isomorfismo de orden, conforme a la terminología a ser introducida en el párrafo siguiente).



Fig. 41

Ejemplo 3. El alfabeto que usamos consiste de las letras

a, b, c, ch, d, e, f, g, h, i, j, k, l, ll,

m, n, ñ, o, p, q, r, s, t, u, v, w, x, y, z

enumeradas en este orden, que forman un conjunto ordenado A de 29 elementos. Daremos el nombre de diccionario al conjunto infinito D de las palabras que se pueden formar con dichas letras. Cada palabra es una secuencia horizontal de un número finito de letras, siendo posible la repetición de las letras. Supongamos que cada palabra comience del lado izquierdo, de manera que la posición de una letra en cada palabra se cuenta a partir de la izquierda. Por ejemplo, *b* ocupa la tercera posición en la palabra *sobremesa*. Entre las diversas palabras, admitimos las que tienen sentido en español, como *miel*, o en portugués, como *mel*, o en inglés, como *honey*, etc., además de las que no tienen sentido en ninguna lengua que nos es conocida, como *crucru*. Cada libro-diccionario de una lengua conocida contiene un número finito de palabras, enumeradas según el orden alfabético, u orden lexicográfico. En realidad, la regla para ordenar las palabras en un libro-diccionario de cualquier lengua es siempre la misma. Ella proviene de una regla de ordenación de D, que se define del siguiente modo: Dadas dos palabras simbolizadas por p_1 y p_2, se dice que p_1 precede a p_2 y se escribe $p_1 \leq p_2$ cuando $p_1 = p_2$, o si $p_1 \neq p_2$, se verifica uno de los dos casos siguientes: 1) el número de letras de p_1 es menor que el número de letras de p_2 y toda letra de p_1 es idéntica a la letra de p_2 en la misma posición, como, por ejemplo, en p_1 = *azúcar*, p_2 = *azucarado*; 2) hay una letra de p_1 que es distinta a la letra de p_2 en la misma posición, tal que la precede en el alfabeto A, y además de esto cada una de las letras anteriores de p_1 coincide

con cada una de las letras anteriores de p_2 en la misma posición, como, por ejemplo, en p_1 = *confitería*, p_2 = *confites*. Es fácil verificar que las cuatro propiedades indicadas en el ejemplo 1 se mantienen válidas en este caso. Para una descripción más formal de este ejemplo, véase el ejercicio 2 que sigue.

Un conjunto *totalmente ordenado*, también llamado cadena, es un conjunto E donde está dada una relación binaria, esto es, una relación entre dos elementos cualesquiera x e y de E, donde se escribe $x \leq y$ para indicar que x e y guardan entre sí la relación considerada, de modo que se satisfacen las siguientes propiedades:

O1. $$x \leq x,$$

O2. $$x \leq y,\ y \leq z \Rightarrow x \leq z,$$

O3. $$x \leq y,\ y \leq x \Rightarrow x = y,$$

O4. $$x \leq y \text{ o } y \leq x,$$

que son las propiedades reflexiva, transitiva, antisimétrica y total, respectivamente. Tal relación binaria recibe el nombre de *orden total*. En circunstancias especiales se usan otros símbolos y otras denominaciones para designar un orden total. Los ejemplos que se acaban de mencionar ilustran el concepto de orden total.

Ejercicios

1) En un conjunto finito E con n elementos existen $n!$ modos distintos de definir un orden total en E.

2) Consideremos un conjunto totalmente ordenado A al cual denominaremos alfabeto. Formemos las potencias cartesianas sucesivas,

$$A^1 = A,\ A^2 = A \times A,\ A^3 = A \times A \times A, \ldots$$

y su unión

$$D = A^1 \cup A^2 \cup \ldots \cup A^n \cup \ldots,$$

a la cual denominaremos diccionario correspondiente al alfabeto A. Cada elemento de D se llama palabra. Si p_1 y p_2 pertenecen a D, escribamos $p_1 \leq p_2$ cuando $p_1 = p_2$, o si $p_1 \neq p_2$, se verifica uno de los dos casos siguientes: 1) el número de coordenadas de p_1 es menor que el número de coordenadas de p_2, y toda coordenada de p_1 es idéntica a la coordenada de p_2 en la misma posición; 2) hay una coordenada de p_1 que es distinta a la coordenada de p_2 en la misma posición y tal que la precede según el orden total de A y, además de esto, cada una de las coordenadas anteriores de p_1 coincide con cada una de las coordenadas anteriores de p_2 en la misma posición. En ambas condiciones 1) y 2) la posición de una coordenada de cada punto de cualquier A^n se cuenta de izquierda a derecha. Entonces D es un conjunto totalmente ordenado.

§ 2. ORDEN

En el caso de un orden total en un conjunto E, la propiedad de que dos elementos arbitrarios x, $y \in E$ siempre sean *comparables*, o sea $x \leq y$ o $y \leq x$, es demasiado exigente. En matemática elemental, como veremos en seguida, ocurren naturalmente relaciones binarias que poseen las tres primeras propiedades de un orden total, pero que dejan de tener la última de ellas. Veamos algunos ejemplos ilustrativos de tal observación.

Ejemplo 1. Consideremos el conjunto $\mathbf{N}$ de los números enteros naturales con la relación de divisibilidad, o sea si x, $y \in \mathbf{N}$, se dice que x divide a y, y se escribe $x|y$, cuando existe $t \in \mathbf{N}$ tal que $y = tx$. Es claro que

$$x|x,$$

$$x|y,\ y|z \Rightarrow x|z,$$

$$x|y,\ y|x \Rightarrow x = y,$$

más no es cierto que

$$x|y \text{ o } y|x,$$

pues, por ejemplo, $2|3$ y $3|2$ son ambas falsas.

Ejemplo 2. En el conjunto $\mathcal{P}(E)$ de todas las partes de un conjunto E tenemos la relación de inclusión y, como ya sabemos,

$$X \subset X,$$

$$X \subset Y,\ Y \subset Z \Rightarrow X \subset Z,$$

$$X \subset Y,\ Y \subset X \Rightarrow X = Y,$$

pero en el caso en que E tiene por lo menos dos elementos distintos no es cierto que

$$X \subset Y \text{ o } Y \subset X,$$

como ya hemos observado.

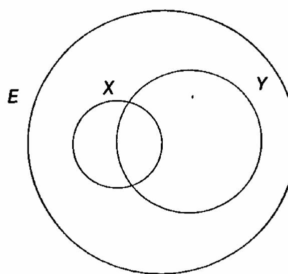


Fig. 42

Ejemplo 3. Consideremos el conjunto F de las funciones reales de variable real definidas en un intervalo $[a, b]$ de $\mathbf{R}$. Si $f, g \in F$ digamos que f es menor o igual que g y escribamos $f \leq g$ cuando $f(x) \leq g(x)$ para todo $x \in [a, b]$, o sea cuando el gráfico de f esté por debajo del gráfico de g.

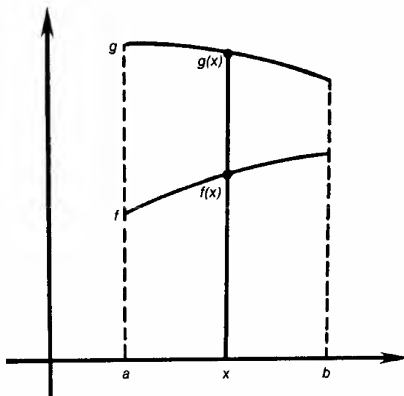


Fig. 43

Es claro que

$$f \leq f,$$

$$f \leq g, \; g \leq h \Rightarrow f \leq h,$$

$$f \leq g, \; g \leq f \Rightarrow f = g,$$

pero si $a \neq b$, no es cierto que

$$f \leq g \text{ o } g \leq f,$$

como se constata si definimos, por ejemplo, $f(x) = x - a$ y $g(x) = b - x$ para todo $x \in [a, b]$.

Un *conjunto ordenado* es un conjunto E donde está dada una relación binaria, esto es una relación entre dos elementos cualesquiera x e y de E, donde se escribe $x \leq y$ para indicar que x e y guardan entre sí la relación considerada, tal que las propiedades 01, 02, 03 del párrafo anterior se satisfacen, más no necesariamente la propiedad 04. Tal relación binaria recibe el nombre de *orden*, u *orden parcial*. En circunstancias especiales se usan otros símbolos y diversas denominaciones para designar un orden. Los ejemplos vistos arriba motivan la introducción del concepto de orden. Todo conjunto totalmente ordenado es un conjunto ordenado, más no recíprocamente.

En un conjunto ordenado E se escribe $x < y$ para indicar que $x \leq y$, pero $x \neq y$. Por tanto, $x \leq y$ si, y sólo si, $x = y$ o $x < y$.

Dado un subconjunto F de un conjunto ordenado E, podemos introducir en F una relación de orden, que se dice *inducida* por la relación

de orden de E, si escribimos $x \leq y$ en F, para x, $y \in F$, si, y sólo si, $x \leq y$ en E. Se dice, entonces, que F es un subconjunto ordenado de E. Por ejemplo, con las relaciones de orden usuales, cada uno de los conjuntos **N**, **Z**, **Q** y **R** es un subconjunto ordenado de los siguientes. Por otra parte, si consideramos a **N** como un conjunto ordenado por la relación de divisibilidad (ejemplo 1), entonces **N** ya no es un subconjunto ordenado de los siguientes con la relación de orden usual. En efecto, basta notar que $2 \leq 3$, pero es falso que $2|3$; por otra parte, $1|0$, pero es falso que $1 \leq 0$.

Dado un conjunto ordenado E, un *elemento mínimo*, o *primer elemento*, de una parte F de E es un elemento $a \in F$ tal que $a \leq x$ para cualquiera que sea $x \in F$. Un tal elemento mínimo a es único, siempre que él exista. En efecto, sea $a' \in F$ tal que también se tenga $a' \leq x$ para todo $x \in F$. Ahora, como $a' \in F$ y del hecho de que a es un elemento mínimo de F resulta que $a \leq a'$. Análogamente, se obtiene $a' \leq a$. Luego $a = a'$. El elemento mínimo de F, cuando existe, se indica por mín F. De un modo dual se define el concepto de *elemento máximo*, o *último elemento*, que se indica por máx F. Por ejemplo, con respecto a sus relaciones de orden usuales, **N** tiene un primer elemento, a saber 0, pero no tiene un último elemento; por otra parte, **Z**, **Q** y **R** no tienen primer ni último elemento. Como otro ejemplo, citemos **N** con respecto a la relación de divisibilidad, donde hay un primer elemento, a saber 1, y también hay un último elemento, esta vez 0. En el caso de $\mathcal{P}(E)$ ordenado por la relación de inclusión, existe un primer elemento $\emptyset$ y un último elemento E.

Sean E y F dos conjuntos ordenados. Se dice que una función $f : E \to F$ es *creciente* cuando $x \leq y$ en E siempre implica $f(x) \leq f(y)$ en F. Por otra parte, f es *decreciente* cuando $x \leq y$ en E siempre implica $f(y) \leq f(x)$ en F. Si f es creciente o decreciente, se dice que f es *monótona*. Decimos que f es estrictamente creciente cuando $x < y$ en E implica siempre $f(x) < f(y)$ en F. Por otra parte, f es *estrictamente decreciente* cuando $x < y$ en E implica siempre $f(y) < f(x)$ en F. Si f es estrictamente creciente o decreciente se dice que f es *estrictamente monótona*.

Ejemplo 4. Sea f una función real de variable real definida y derivable en un intervalo $[a, b]$ de R. Si $f'(x) \geq 0$ para todo $x \in [a, b]$, entonces f es creciente. Si $f'(x) > 0$ para todo $x \in [a, b]$, entonces f es estrictamente creciente. Si $f'(x) \leq 0$ para todo $x \in [a, b]$, entonces f es decreciente. Si $f'(x) < 0$ para todo $x \in [a, b]$, entonces f es estrictamente decreciente.

Una función creciente a veces se llama *no decreciente* y en tal caso se entiende por función *creciente* lo que llamamos función estrictamente creciente. Es claro que, entonces, se acostumbra decir función *no creciente* en vez de función decreciente, así como función *decreciente* para designar lo que llamamos función estrictamente decreciente.

Ejemplo 5. Consideremos el conjunto C de las funciones reales continuas en el intervalo $[a, b]$. En C, empleemos la relación de orden inducida por el conjunto ordenado F de las funciones reales en $[a, b]$ (véase el ejemplo 3). Entonces

$$f \in C \mapsto \int_a^b f(x)dx \in \mathbf{R}$$

es estrictamente creciente.

Ejemplo 6. Dada la función $f : E \to F$, podemos considerar la función

$$X \in \mathcal{P}(E) \mapsto f(X) \in \mathcal{P}(F),$$

que es siempre creciente. Para que esta última sea estrictamente creciente, es necesario y suficiente que la función dada sea biunívoca. Análogamente, podemos considerar la función

$$Y \in \mathcal{P}(F) \mapsto f^{-1}(Y) \in \mathcal{P}(E),$$

que es siempre creciente. Para que esta última sea estrictamente creciente, es necesario y suficiente que la función dada sea sobre.

Dados dos conjuntos ordenados E y F, llámase *isomorfismo* entre E y F a toda aplicación biunívoca f de E sobre F tal que $x \leq y$ en E si, y sólo si, $f(x) \leq f(y)$ en F. Un isomorfismo entre E y F es, pues, una aplicación biunívoca f de E sobre F tal que f y f^{-1} sean crecientes (o, equivalentemente, que sean estrictamente crecientes). E y F son *isomorfos* cuando existe un tal isomorfismo. Un *antiisomorfismo* entre E y F es toda aplicación biunívoca f de E sobre F tal que $x \leq y$ en E si, y sólo si, $f(y) \leq f(x)$ en F. Con tal fin, se hacen comentarios análogos a los que se hicieron respecto a un isomorfismo.

Ejemplo 7. La función $x \mapsto e^x$ es un isomorfismo de orden entre $\mathbf{R}$ y $\mathbf{R}_+^*$ Su isomorfismo inverso es la función $x \mapsto \log x$ entre $\mathbf{R}_+^*$ y $\mathbf{R}$. El mismo tipo de ejemplo se repite para toda función real f de variable real que sea estrictamente creciente entre su dominio $E \subset \mathbf{R}$ y su contradominio $f(E) \subset \mathbf{R}$, considerándose a E y a $f(E)$ ordenados por el orden inducido por $\mathbf{R}$. En el caso de ser f estrictamente decreciente, se tiene un antiisomorfismo de orden entre E y $f(E)$.

Ejemplo 8. Sea $\mathbf{G}(\mathbf{Z})$ el conjunto de los subgrupos del grupo aditivo $\mathbf{Z}$ de los enteros racionales. A todo $H \in \mathbf{G}(\mathbf{Z})$ (o sea, a todo subgrupo H de $\mathbf{Z}$) le corresponde uno, y sólo un, entero natural $n(H) \in \mathbf{N}$ tal que H es el conjunto de los múltiplos enteros de $n(H)$. La función $H \in \mathbf{G}(\mathbf{Z}) \mapsto n(H) \in \mathbf{N}$ es un antiisomorfismo de orden entre $\mathbf{G}(\mathbf{Z})$ ordenado por la relación de inclusión y $\mathbf{N}$ ordenado por la relación de divisibilidad De modo análogo, dado un entero $p \geq 1$, sea $\mathbf{G}(\mathbf{Z}/p)$ el conjunto de los subgrupos del grupo aditivo $\mathbf{Z}/p$ de los enteros $0, 1, \ldots, p-1$ módulo p. Indiquemos con $D(p)$ al conjunto de los enteros naturales que dividen a p. A todo $H \in \mathbf{G}(\mathbf{Z}/p)$ (es decir, a todo subgrupo H de $\mathbf{Z}/p$) le corresponde uno, y sólo un, entero natural $n(H) \in D(p)$ tal que H es el conjunto de los múltiplos enteros naturales de $n(H)$ que pertenecen a $\mathbf{Z}/p$. La función $H \in \mathbf{G}(\mathbf{Z}/p) \mapsto n(H) \in D(p)$ es un antiisomorfismo de orden entre $\mathbf{G}(\mathbf{Z}/p)$ ordenado por la relación de inclusión y $D(p)$ ordenado por la relación de divisibilidad. De ahí resulta que $\mathbf{G}(\mathbf{Z}/p)$ es totalmente ordenado si, y sólo si, $D(p)$ es totalmente ordenado, lo que equivale a que p sea una potencia entera de un entero natural primo.

Ejemplo 9. La función

$$X \in \mathcal{P}(E) \mapsto \complement X \in \mathcal{P}(E)$$

es un antiisomorfismo de orden.

Ejemplo 10. Consideremos el conjunto $\mathbf{N}^*$ de los números enteros naturales diferentes de 0 ordenados por la relación de divisibilidad. Por otra parte, sea S el conjunto de las sucesiones

$$(n_1, n_2, \dots, n_k, \dots)$$

de números enteros naturales, o sea $n_k \in \mathbf{N}$ para $k = 1, 2, \dots,$ y asumamos también que cada una de las sucesiones sólo contiene un número finito de términos distintos de cero, o sea $n_k = 0$ excepto para un número finito de valores de k. Definamos un orden en S por

$$(m_1, m_2, \dots, m_k, \dots) \leq (n_1, n_2, \dots, n_k, \dots)$$

cuando $m_1 \leq n_1, m_2 \leq n_2, \dots, m_k \leq n_k, \dots$. Sea ahora

$$p_1 = 2, \ p_2 = 3, \ p_3 = 5, \ p_4 = 7, \dots$$

la sucesión de los números enteros naturales primos escrita en el mismo orden en que se presenta en $\mathbf{N}$. A cada sucesión

$$(n_1, n_2, \dots, n_k, \dots) \in S$$

podemos asociar el entero

$$(p_1)^{n_1} (p_2)^{n_2} \dots (p_k)^{n_k} \in \mathbf{N}^*,$$

donde k se eligió suficientemente grande tal que $n_r = 0$ para todo $r > k$; notemos que la elección de un tal k no afecta el valor del elemento de $\mathbf{N}^*$ así definido. Acabamos, pues, de definir una función $f : S \to \mathbf{N}^*$. El teorema básico de aritmética elemental sobre la unicidad y la existencia de la descomposición de todo elemento de $\mathbf{N}^*$ en el producto de potencias enteras naturales de factores primos afirma precisamente que f es biunívoca en S y sobre $\mathbf{N}^*$. Por otra parte, el teorema básico de aritmética elemental sobre la divisibilidad de un elemento por otro de $\mathbf{N}^*$, divisibilidad ésta expresada en términos de las descomposiciones de estos dos números en productos de potencias enteras naturales de factores primos, afirma precisamente que f es un isomorfismo de orden entre S y $\mathbf{N}^*$.

En el conjunto ordenado E, dos elementos x e y se dicen *consecutivos* cuando $x < y$, y además no existe ningún t tal que $x < t < y$. Esta noción permite ilustrar la descripción de todo conjunto finito por medio de su diagrama: éste se obtiene por medio de figuras rectilíneas o del plano (y, en algunos casos, inclusive por figuras imaginadas en el espacio tridimensional), uniendo a los pares de puntos consecutivos de E por segmentos o arcos orientados de cada punto que precede a cada punto que sigue (véase el ejercicio 1 más abajo).

Un ejemplo ilustrativo es el del diagrama del conjunto de los enteros naturales que dividen a 12 ordenados según la relación de divisibilidad.

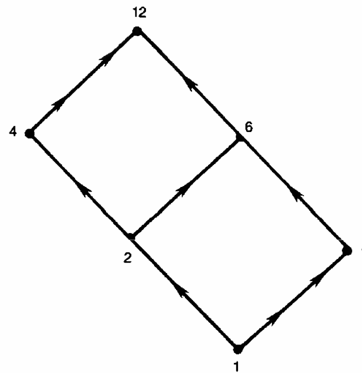


Fig. 44

Otro ejemplo aclaratorio es el del diagrama del conjunto $\mathcal{P}(E)$ de las partes de un conjunto $E = \{x_1, x_2, x_3\}$ formado por tres puntos. Sabemos que $\mathcal{P}(E)$ queda ordenado por la relación de inclusión y consta de $2^3 = 8$ elementos, a saber

$$\phi,$$

$$E_1 = \{x_1\},\ E_2 = \{x_2\},\ E_3 = \{x_3\},$$

$$E_{12} = \{x_1, x_2\},\ E_{13} = \{x_1, x_3\},\ E_{23} = \{x_2, x_3\},$$

$$E.$$

El diagrama $\mathcal{P}(E)$ debe imaginarse formado por los ocho vértices de un cubo en el espacio tridimensional.

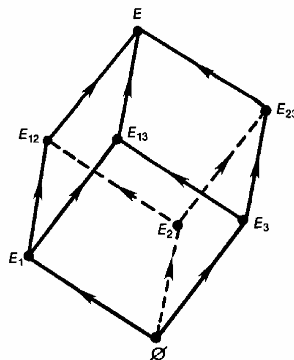


Fig. 45

Ejercicios

1) Supongamos que un conjunto ordenado E satisface la condición siguiente: si x, $y \in E$ con $x \leq y$, entonces el conjunto de los elementos $t \in E$ tales que $x \leq t \leq y$ es finito (esta condición se verifica cuando E es finito). Probar que dados x, $y \in E$ con $x < y$, existen $n \geq 1$ elementos $t_0, t_1, \ldots, t_n \in E$ tales que $t_0 = x$, $t_n = y$ y t_{i-1}, t_i son consecutivos para $i = 1, \ldots, n$. Concluir que el conocimiento de los pares ordenados (x, y) de elementos consecutivos en E, con $x < y$, determina el conocimiento del orden en E.

2) Sea E un conjunto ordenado. El gráfico G del orden de E es el subconjunto del cuadrado cartesiano E^2 formado por los pares (x, y), donde x, $y \in E$ y $x \leq y$. Demostrar que

$$G \cap G^{-1} = \Delta, \; GG \subset G$$

(notación de los ejercicios 3 y 4 del párrafo 9, capítulo 1). Recíprocamente, dada una parte G de E^2 que verifique las dos condiciones de arriba, con E un conjunto, si definimos $x \leq y$ cuando $(x, y) \in G$, obtenemos una relación de orden en E cuyo gráfico es G. Para que se tenga un orden total es necesario y suficiente que

$$G \cup G^{-1} = E^2.$$

3) Sea E un conjunto ordenado. El gráfico estricto G del orden de E es el subconjunto de E^2 formado por los pares (x, y), donde x, $y \in E$ y $x < y$. Demostrar que

$$G \cap G^{-1} = \phi, \; GG \subset G.$$

Recíprocamente, dada una parte G de E^2 que verifique las dos condiciones de arriba, con E un conjunto, si definimos $x \leq y$ cuando $x = y \in E$ o cuando $(x, y) \in G$, obtenemos una relación de orden en E cuyo gráfico estricto es G. Para que se tenga un orden total es necesario y suficiente que

$$\Delta \cup G \cup G^{-1} = E^2.$$

4) Si $f : E \to F$ es biunívoca en E, sobre F y creciente, donde E es un conjunto totalmente ordenado y F es un conjunto ordenado, entonces F es totalmente ordenado y f es un isomorfismo de orden entre E y F.

5) Sean $f : E \to F$ y $g : F \to G$ dos funciones, donde E, F y G son conjuntos ordenados. Con relación a la función compuesta $gf : E \to G$ mostrar que gf es creciente cuando f y g son ambas crecientes, o ambas decrecientes, y que gf es decreciente cuando una de f o g es creciente y la otra es decreciente.

6) Un conjunto bien ordenado es un conjunto ordenado E, donde toda parte no vacía tiene un primer elemento. Todo conjunto bien ordenado es totalmente ordenado. Todo conjunto totalmente ordenado finito es bien ordenado. $\mathbf{N}$ es bien ordenado. El subconjunto de R de los números racionales de la forma

$$m - \frac{1}{n} \; (m, \;\; n = 1, \; 2, \ldots)$$

es bien ordenado.

7) Un conjunto preordenado es un conjunto E donde está dada una relación de preorden, que se indica por $x \leq y$ para $x, \; y \in E$, de modo que

$$x \leq x,$$

$$x \leq y, \;\; y \leq z \Rightarrow x \leq z.$$

Si definimos $x \sim y$ cuando $x \leq y, \; y \leq x$, donde $x, \; y \in E$, obtenemos una relación de equivalencia en E. Sea F el espacio cociente de E por esta relación. Si $\overline{x}, \; \overline{y} \in F$ son las clases de equivalencia correspondientes a $x, \; y \in E$, definamos $\overline{x} \leq \overline{y}$ en F cuando $x \leq y$ en E. Probar que F es un conjunto ordenado.

8) Si A fuese un anillo con unidad y $x, \; y \in A$, se dice que x divide a y a la derecha cuando existe un $t \in A$ tal que $y = tx$. Probar que, de tal modo, se obtiene un preorden en A. Análogamente, para la divisibilidad a la izquierda.

§ 3. RETICULADOS

Dada una parte X de un conjunto ordenado E, se dice que X está *limitado superiormente* en E cuando existe un elemento $s \in E$, tal que $x \leq s$ para todo $x \in X$. A un tal elemento s se le denomina *mayorante* o *cota superior* de X en E. Notemos que una cota superior puede no ser única. Si X no está limitado superiormente en E se dice que X es *ilimitado superiormente* en E. Observemos que si E tuviera un último elemento, entonces todo subconjunto de E estará limitado superiormente en E y que, recíprocamente, si E estuviera limitado superiormente en E, entonces E deberá poseer un último elemento.

Admitiendo que X esté limitado superiormente en E, puede ocurrir que el conjunto no vacío de las cotas superiores de X tenga un elemento mínimo, esto es que exista una cota superior s de X en E tal que cualquier otra cota superior de X en E deba satisfacer $s \leq t$. En tal caso, el elemento mínimo de las cotas superiores de X en E (que es único) se denomina *supremo* de X en E y se representa por sup X (u otras notaciones, dependiendo del caso). Si X tuviera un máximo, es claro que X tendrá un supremo en E y que

$$\text{máx}\; X = \sup X.$$

De un modo dual, se definen los conceptos de subconjunto limitado inferiormente, de minorante o cota inferior, de subconjunto ilimitado inferiormente y de ínfimo (ínf X) de X en E y se hacen los comentarios duales pertinentes.

Se dice que un conjunto ordenado E es un *reticulado* cuando, cualesquiera que sean los elementos $x, y \in E$, el conjunto $\{x, y\}$ reducido a los mismos tiene un supremo

$$\sup \{x, y\},$$

que se indica también por

$$x \cup y,$$

así como un ínfimo

$$\inf \{x, y\},$$

que se representa igualmente por

$$x \cap y.$$

Todo conjunto totalmente ordenado es un reticulado, pues

$$x \cup y = y, \; x \cap y = x, \text{ si es que } x \leq y.$$

Ejemplo 1. Consideremos otra vez el conjunto **N** de los enteros naturales ordenado por la relación de divisibilidad (ejemplo 1 de la sección precedente). Dados $x, y \in \mathbf{N}$, el supremo y el ínfimo de x e y en **N** existen necesariamente y son los que en aritmética elemental se llaman el mínimo común múltiplo y el máximo común divisor de x e y. Luego, **N** es un reticulado con respecto a la divisibilidad.

Ejemplo 2. El conjunto $\mathcal{P}(E)$ de todas las partes de un conjunto E, ordenado por la relación de inclusión, es un reticulado, donde el supremo y el ínfimo de dos partes de E son precisamente su unión y su intersección.

Ejemplo 3. El conjunto F de todas las funciones reales de una variable real definidas en un intervalo $\{a, b\}$ de **R** es un reticulado con

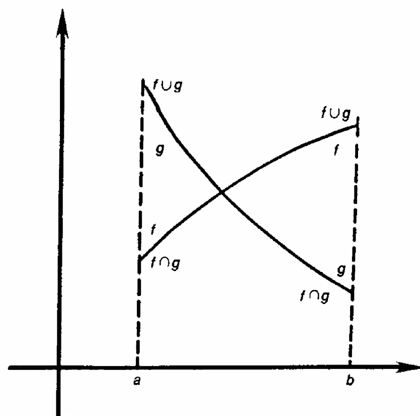


Fig. 46

respecto al orden usual (ejemplo 3 de la sección precedente). Si f, $g \in F$, entonces su supremo y su ínfimo en F se definen, respectivamente, por

$$(f \cup g)(x) = \sup \{f(x), g(x)\},$$

$$(f \cap g)(x) = \inf \{f(x), g(x)\}.$$

Proposición 1. Para elementos arbitrarios x, y, z de un reticulado E, se tiene:

1) $x \cup y = y \cup x$, $\qquad x \cap y = y \cap x$,

2) $x \cup (y \cup z) = (x \cup y) \cup z$, $\qquad x \cap (y \cap z) = (x \cap y) \cap z$,

3) $y \leq x \Leftrightarrow x \cup y = x$, $\qquad x \leq y \Leftrightarrow x \cap y = x$.

La demostración se deja como ejercicio.

Cualesquiera que sean los elementos $x_1, x_2, \ldots, x_n$ en número finito n, no nulo, de un reticulado E, el conjunto finito $\{x_1, x_2, \ldots, x_n\}$ reducido a los mismos tiene un supremo en E

$$\sup \{x_1, x_2, \ldots, x_n\},$$

que se indica también por

$$x_1 \cup x_2 \cup \ldots \cup x_n,$$

así como un ínfimo

$$\text{ínf} \{x_1, x_2, \ldots, x_n\},$$

que se indica también por

$$x_1 \cap x_2 \cap \ldots \cap x_n,$$

y basta proceder por inducción sobre n para concluir que el supremo y el ínfimo indicados existen en E. A propósito, observemos que

$$x \cup (y \cup z) = x \cup y \cup z,$$

$$(x \cup y) \cup z = x \cup y \cup z;$$

análogamente

$$x \cap (y \cap z) = x \cap y \cap z,$$

$$(x \cap y) \cap z = x \cap y \cap z.$$

Se dice que un subconjunto F de un reticulado E es un *subreticulado* de E cuando F es un reticulado con respecto al orden inducido en F por E y cuando el supremo y el ínfimo de dos elementos x, $y \in F$ tienen los mismos valores, sea que se calculen considerando a x e y como ele-

mentos de E o como elementos de F. Todo subconjunto de un conjunto totalmente ordenado es un subreticulado.

Proposición 2. Para que el subconjunto F del reticulado E sea un subreticulado de E es necesario y suficiente que $x, y \in F$ impliquen $x \cup y \in F$ y $x \cap y \in F$, donde el supremo y el ínfimo en cuestión son calculados en E.

La demostración se deja como ejercicio.

Ejemplo 4. Consideremos el reticulado F de las funciones reales de variable real definidas en el intervalo $\{a, b\}$ de $\mathbf{R}$ (ejemplo 3). Afirmamos que el conjunto C de las funciones reales continuas en $\{a, b\}$ es un subreticulado de F. En efecto, como se sabe, el supremo y el ínfimo de dos tales funciones continuas son también funciones continuas.

Se dice que un reticulado E es *distributivo* cuando

$$x \cap (y \cup z) = (x \cap y) \cup (x \cap z),$$

$$x \cup (y \cap z) = (x \cup y) \cap (x \cup z),$$

cualesquiera que sean $x, y, z \in E$. Los ejemplos 1, 2 y 3 de reticulados son todos distributivos. En el caso del ejemplo 1, se observa en primer lugar que $\mathbf{N}^*$ es un subreticulado de $\mathbf{N}$ y que $\mathbf{N}^*$ es un reticulado distributivo, pues (ejemplo 10 de la sección prededente) hay un isomorfismo de orden entre $\mathbf{N}^*$ y S, el cual es, a las claras, un reticulado distributivo. En el caso del ejemplo 2, ya se sabe (proposición 1, §3, capítulo 1) que $\mathcal{P}(E)$ es distributivo. En cuanto al ejemplo 3, la verificación de su distributividad es inmediata.

Ejemplo 5. Es fácil verificar que los conjuntos ordenados de cinco elementos, representados por los diagramas de la figura 47, son ambos reticulados no distributivos. Por tanto, todo reticulado que contenga un subreticulado isomorfo, en el sentido del orden, a uno de tales reticulados de cinco elementos no puede ser distributivo. A este propósito, se puede probar recíprocamente que todo reticulado no distributivo contiene siempre un subreticulado que es isomorfo, en el sentido del orden, a uno de tales reticulados de cinco elementos.

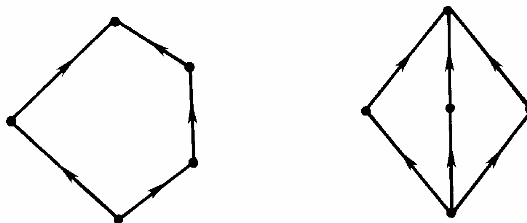


Fig. 47

Sean E y F dos reticulados. Se dice que una función $f : E \to F$ es un *homomorfismo* de reticulados cuando

$$f(x \cup y) = f(x) \cup f(y),$$

$$f(x \cap y) = f(x) \cap f(y),$$

cualesquiera que sean $x, y \in E$. Todo homomorfismo de reticulados es creciente, pues si $x, y \in E$ y $x \leq y$, entonces $x \cup y = y$, de donde $f(x) \cup f(y) = f(y)$, o sea $f(x) \leq f(y)$. Todo homomorfismo de reticulados f que sea biunívoco en E y sobre F es necesariamente un isomorfismo de orden entre E y F. Análogamente, se define un *antihomomorfismo* de reticulados por medio de las condiciones

$$f(x \cup y) = f(x) \cap f(y),$$

$$f(x \cap y) = f(x) \cup f(y).$$

Ejemplo 6. Dada la función $f : E \to F$, la función

$$Y \in \mathcal{P}(F) \mapsto f^{-1}(Y) \in \mathcal{P}(E)$$

es un homomorfismo de reticulados; no ocurre lo mismo con la función

$$X \in \mathcal{P}(E) \mapsto f(X) \in \mathcal{P}(F),$$

a menos que la función dada sea biunívoca.

Ejercicios

1) En un conjunto ordenado E, se define el intervalo $[x, y]$ (cuando $x, y \in E$ y $x \leq y$) como los $t \in E$ tales que $x \leq t \leq y$. Si F es un reticulado, entonces la intersección de dos intervalos de F o es un intervalo o es vacía. Recíprocamente, sea E un conjunto ordenado donde la intersección de dos intervalos en E es un intervalo o es vacía y donde dos elementos cualesquiera de E siempre tienen una cota inferior y una cota superior. Entonces, E es un reticulado.

2) Sea E un conjunto y supongamos dadas dos funciones

$$(x, y) \in E^2 \mapsto x \cup y \in E,$$

$$(x, y) \in E^2 \mapsto x \cap y \in E,$$

tales que

$$x \cap (y \cup z) = (x \cap y) \cup (x \cap z), \qquad x \cup (y \cap z) = (x \cup y) \cap (x \cup z),$$

$$x \cup y = y \cup x, \qquad x \cap y = y \cap x,$$

$$x \cup (y \cup z) = (x \cup y) \cup z, \qquad x \cap (y \cap z) = (x \cap y) \cap z,$$

$$x \cup x = x, \qquad x \cap x = x.$$

Demostrar que $x \cup y = y$ equivale a $x \cap y = x$ y que si definiéramos $x \leq y$ por tales igualdades E resultaría ser un conjunto ordenado que es un reticulado distributivo, donde $x \cup y$ y $x \cap y$ son el supremo y el ínfimo de x e $y \in E$.

3) Sea P un plano y E el conjunto ordenado por inclusión, cuyos elementos son P, las rectas de P, los subconjuntos de P que se reducen a puntos de P y el conjunto vacío. Entonces, E es un reticulado no distributivo y E no es un subreticulado del reticulado de los subconjuntos de P.

4) Dada una función $f : E \to E$, probar que el conjunto de los subconjuntos X de E tales que $f(X) \subset X$ es un subreticulado de $\mathcal{P}(E)$, ordenado por la inclusión. Por otra parte, el conjunto de los subconjuntos X de E tales que $X \subset f(X)$ no es obligatoriamente un subreticulado de $\mathcal{P}(E)$.

§ 4. ALGEBRAS DE BOOLE

Consideremos un reticulado distributivo E con primer elemento 0 y último elemento I. Si $x \in E$, se llama *complemento* de x en E y se representa por

$$\mathbf{C}x$$

 a todo elemento de E tal que

$$x \cap (\mathbf{C}x) = 0, \qquad x \cup (\mathbf{C}x) = I.$$

Mostremos, antes de continuar, que un tal complemento de x es único, siempre que exista. En efecto, si $y,\ z \in E$ y tanto y como z fueran complementos de x en E, esto es si

$$x \cap y = 0, \qquad x \cup y = I,$$

$$x \cap z = 0, \qquad x \cup z = I,$$

entonces tendremos

$$y \cap (x \cup z) = y \cap I,$$

de donde

$$(y \cap x) \cup (y \cap z) = y,$$

o sea

$$0 \cup (y \cap z) = y.$$

Por tanto, $y \cap z = y$, de donde $y \leq z$. Análogamente, se ve que $z \leq y$, de donde $y = z$.

Un *álgebra de Boole* es un reticulado distributivo con primer y último elementos, donde todo elemento tiene un complemento.

Ejemplo 1. A todo conjunto E le corresponde el álgebra de Boole $\mathcal{P}(E)$ de todos los subconjuntos de E, con respecto a la inclusión.

Ejemplo 2. El ejemplo que vamos a presentar se debe a Boole y es el que, históricamente, dió origen a la denominación de álgebra de Boole. Tal ejemplo tiene por finalidad describir las propiedades elementales que admitimos para la lógica de las proposiciones. Pasemos, pues, a describir intuitivamente lo que supondremos a tal respecto. Sean dadas dos proposiciones

$$p, \ q.$$

Tenemos el concepto de que la proposición p implica lógicamente la proposición q, lo que usualmente se escribe

$$p \Rightarrow q.$$

Tenemos, también, el concepto de que p y q son lógicamente equivalentes, o sea

$$p \Rightarrow q \text{ y } q \Rightarrow p,$$

lo que usualmente se escribe

$$p \Leftrightarrow q.$$

Notemos que p y q determinan dos nuevas proposiciones a saber la proposición 139

$$p \text{ o } q$$

y la proposición

$$p \text{ y } q.$$

Finalmente, cada proposición determina su proposición negativa

$$\text{no } p.$$

Supongamos ahora dado un conjunto no vacío A de proposiciones, que satisface las condiciones siguientes:

1) Si $p, \ q \in A$, entonces

$$(p \text{ y } q) \in A, \qquad (p \text{ o } q) \in A.$$

2) Si $p \in A$, entonces

$$(\text{no } p) \in A.$$

Admitiremos que la equivalencia lógica entre las proposiciones de A sea una relación de equivalencia en A. Indiquemos con B el correspondiente conjunto cociente de A. Si $\overline{p}$ y $\overline{q}$ fueran las clases de equivalencia correspondientes a $p, \ q \in A$, escribiremos $\overline{p} \leq \overline{q}$ cuando $p \Rightarrow q$.

Admitiremos, entonces, que B es un conjunto ordenado y, más precisamente, un álgebra de Boole donde

$$\begin{aligned}
\overline{p \text{ o } q} &= \overline{p} \cup \overline{q}, \\
\overline{p \text{ y } q} &= \overline{p} \cap \overline{q}, \\
\overline{p \text{ y (no } p)} &= 0, \\
\overline{p \text{ o (no } p)} &= I, \\
\overline{\text{no } p} &= \mathbf{C}\overline{p}
\end{aligned}$$

Para una descripción más formal de la situación en este ejemplo, véase el ejercicio 2 más abajo.

Una *subálgebra de Boole* F de un álgebra de Boole E es un subreticulado de E que contiene el primer y último elementos de E y que igualmente es un álgebra de Boole. Es inmediato que el primer y el último elementos de E son también primer y último elementos de F, además de que el complemento en F de cualquier elemento de F coincide con su complemento en E.

Ejemplo 3. En una recta R tenemos el concepto de intervalo, que puede ser limitado o ilimitado, así como cerrado o abierto en cada una de sus extremidades, conforme contenga o no a tal extremidad. Sea S el conjunto de los subconjuntos de $\mathbf{R}$ cada uno de los cuales es la unión de un número finito de tales subconjuntos. Entonces, con respecto a la inclusión, S es una subálgebra de Boole de $\mathcal{P}(\mathbf{R})$.

Sean E y F dos álgebras de Boole. Se dice que una función $f : E \to F$ es un *homomorfismo* de álgebras de Boole cuando

$$\begin{aligned}
f(x \cup y) &= f(x) \cup f(y), \\
f(x \cap y) &= f(x) \cap f(y), \\
f(\mathbf{C}x) &= \mathbf{C}f(x)
\end{aligned}$$

para cualesquiera $x,\ y \in E$. Notemos que, entonces,

$$f(0) = 0, \qquad\qquad f(I) = I.$$

En efecto,

$$f(0 \cup \mathbf{C}0) = f(0) \cup \mathbf{C}f(0) = I$$

lo que prueba que $f(I) = I$. Análogamente, se verifica que $f(0) = 0$. El concepto de *antihomomorfismo* se define de modo semejante.

Ejemplo 4. Dada la función $f : E \to F$, la función

$$Y \in \mathcal{P}(F) \mapsto f^{-1}(Y) \in \mathcal{P}(E)$$

es un homomorfismo de álgebras de Boole; no ocurre lo mismo con la función

$$X \in \mathcal{P}(E) \mapsto f(X) \in \mathcal{P}(F),$$

a menos que la función dada f sea biunívoca en E sobre F.

EJERCICIOS

1) Toda álgebra de Boole finita tiene 2^n elementos, siendo isomorfa al álgebra de Boole de los subconjuntos de un conjunto finito de n elementos.

2) Sea E un conjunto preordenado y consideremos la relación de equivalencia correspondiente (ejercicio 7, §2). Supongamos dadas las funciones

$$(x, y) \in E^2 \mapsto x \cup y \in E,$$

$$(x, y) \in E^2 \mapsto x \cap y \in E,$$

$$x \in E \mapsto \complement x \in E$$

y admitamos que

$$x \leq z \text{ e } y \leq z \Leftrightarrow x \cup y \leq z,$$

$$z \leq x \text{ y } z \leq y \Leftrightarrow z \leq x \cup y,$$

$$x \cap \complement x \leq y \leq x \cup \complement x,$$

$$x \cap (y \cup z) \sim (x \cap y) \cup (x \cap z),$$

$$x \cup (y \cap z) \sim (x \cup y) \cap (x \cup z),$$

cualesquiera que sean $x, y, z \in E$. Sea F el conjunto ordenado cociente de E por esa relación de equivalencia. Probar que F es un álgebra de Boole.

COLECCION DE MONOGRAFIAS CIENTIFICAS

Publicadas

Serie de matemática

N° 1. La Revolución en las Matemáticas Escolares, por el Consejo Nacional de Maestros de Matemáticas de los Estados Unidos de América.
N° 2. Espacios Vectoriales y Geometría Analítica, por Luis A. Santaló.
N° 3. Estructuras Algebraicas I, por Enzo R. Gentile.
N° 4. Historia de las Ideas Modernas en la Matemática, por José Babini.
N° 5. Algebra Lineal, por Orlando E. Villamayor.
N° 6. Algebra Linear e Geometría Euclideana, por Alexandre Augusto Martins Rodrigues.
N° 7. El Concepto de Numero, por Cesar A. Trejo.
N° 8. Funciones de Variable Compleja, por José I. Nieto.
N° 9. Introducción a la Topología General, por Juan Horváth.
N° 10. Funções Reais, por Djairo G. de Figueiredo.
N° 11. Probabilidad e Inferencia Estadística, por Luis A. Santaló.
N° 12. Estructuras Algebraicas II (Algebra Lineal), por Enzo R. Gentile.
N° 13. La Revolución en las Matemáticas Escolares (Segunda Fase), por Howard F. Fehr, John Camp y Howard Kellog.
N° 14. Estructuras Algebraicas III (Grupos Finitos), por Horacio H. O'Brien.
N° 15. Introducción a la Teoría de Grafos, por Fausto A. Toranzos.
N° 16. Estructuras Algebraicas IV (Algebra Multilineal), por Artibano, Micali y Orlando E. Villamayor.
N° 17. Introdução a Análise Funcional: Espaços de Banach e Cálculo /Diferencial, por Leopoldo Nachbin. 143
N° 18. Introducción a la Integral de Lebesgue en la Recta, por Juan Antonio Gatica.
N° 19. Introducción a los Espacios de Hilbert, por José I. Nieto.
N° 20. Elementos de Biomatemática, por Alejandro B. Engel.
N° 21. Introducción a la Computación, por Jaime Michelow.
N° 22. Estructuras Algebraicas V (Teoría de Cuerpos), por Héctor A.
a. Merklen.
N° 23. Estructuras Algebraicas VI (Formas Cuadráticas), por /Francisco M. Piscoya.
N° 24. Estructuras Algebraicas VII (Estructuras de Algebras), por Artibano Micali.
N° 25. Aritmética Elemental, por Enzo R. Gentile.
N° 26. Algebra Elemental, por Leopoldo Nachbin.

a

143

epto Moderno del Núcleo, por D. Allan Bromley.
rama de la Astronomía Moderna, por Félix Cernuschi y Sayd Codina.
structura Electrónica de los Sólidos, por Leopoldo M. Falicov.

N° 4. Física de Partículas, por Igor Saavedra.
N° 5. Experimento, Razonamiento y Creación en Física, por Félix Cernuschi.
N° 6. Semiconductores, por George Bemski.
N° 7. Aceleradores de Partículas, por Fernando Alba Andrade.
N° 8. Física Cuántica, por Onofre Rojo y Harold V. McIntosh.
N° 9. La Radiación Cósmica, por Gastón R. Mejía y Carlos Aguirre.
N° 10. Astrofísica, por Carlos Jaschek y Mercedes C. de Jaschek.
N° 11. Ondas, por Oscar J. Bressan y Enrique Gaviola.
N° 12. El Láser, por Mario Garavaglia.
N° 13. Teoría Estadística de la Materia, por Antonio E. Rodríguez y Roberto E. Caligaris.
N° 14. Aplicações da Teoria de Grupos na Espectroscopia Raman e do Infra-Vermelho, por Jorge Humberto Nicola y Anildo Bristoti.
N° 15. Fundamentos de Cristalografía Física, por Jaime Rodríguez Lara.

Serie de química

N° 1. Cinética Química Elemental, por Harold Behrens Le Bas.
N° 2. Bioenergética, por Isaias Raw y Walter Colli.
N° 3. Macromoléculas, por Alejandro Paladini y Moisés Burachik.
N° 4. Mecanismo de las Reacciones Orgánicas, por Jorge A. Brieux.
N° 5. Elementos Encadenados, por Jacobo Gómez Lara.
N° 6. Enseñanza de la Química Experimental, por Francisco Giral.
N° 7. Fotoquímica de Gases, por Ralf-Dieter Penzhorn.
N° 8. Introducción a la Geoquímica, por Félix González-Bonorino.
N° 9. Resonancia Magnética Nuclear de Hidrógeno-1 y de Carbono-13, por Pedro Joseph-Nathan.
N° 10. Cromatografía Líquida de Alta Presión, por Harold M. McNair y Benjamín Esquivel H.
N° 11. Actividad Optica, Dispersión Rotatoria Optica y Dicroísmo Circular en Química Orgánica, por Pierre Crabbé.
N° 12. Espectroscopia Infrarroja, por Jesús Morcillo Rubio.
N° 13. Polarografía, por Alejandro J. Arvía y Jorge A. Bolzán.
N° 14. Paramagnetismo Electrónico, por Juan A. McMillan.
N° 15. Introducción a la Estereoquímica, por Juan A. Garbarino.
N° 16. Cromatografía en Papel y en Capa Delgada, por Xorge A. Domínguez.
N° 17. Introducción a la Espectrometría de Masa de Sustancias Orgánicas, por Otto R. Gottlieb y Raimundo Braz Filho.
N° 18. Cinética Química, por Rodolfo V. Caneda.
N° 19. Fuerzas Intermoleculares, por Mateo Díaz Peña.
N° 20. Físico-Química de Superficies, por Tibor Rabockai.
N° 21. Corrosión, por José R. Galvele.
N° 22. Introducción a la Electroquímica, por Dionisio Posadas.
N° 23. Cromatografía de Gases, por Harold M. McNair.
N° 24. Cinética de Disolución de Medicamentos, por Edison Cid Cárcamo.
N° 25. Introducción a la Química de Suelos, por Elemer Bornemisza.
N° 26. Elementos de Catálisis Heterogénea, por Sergio E. Droguett.
N° 27. Introducción a la Electrocatálisis, por Alejandro J. Arvía y María Cristina Giordano.

N° 28. Química de Sólidos, por Julio César Bazán.
N° 29. Química Bioinorgânica, por Henrique Eisi Toma.
N° 30. Introducción al Estudio de los Productos Naturales, por Eduardo G. Gros, Alicia B. Pomilio, Alicia M. Seldes y Gerardo Burton.

Serie de biología

N° 1. La Genética y la Revolución en las Ciencias Biológicas, por José Luis Reissig.
N° 2. Bases Ecológicas para la Explotación Agropecuaria en la América Latina, por Guillermo Mann F.
N° 3. La Taxonomía y la Revolución en las Ciencias Biológicas, por Elías R. de la Sota.
N° 4. Principios Básicos para la Enseñanza de la Biología, por Oswaldo Frota-Pessoa.
N° 5. A Vida da Célula, por Renato Basile.
N° 6. Microorganismos, por J. M. Gutiérrez-Vázquez.
N° 7. Principios Generales de Microbiología, por Norberto J. Palleroni.
N° 8. Los Virus, por Enriqueta Pizarro-Suárez y Gamba.
N° 9. Introducción a la Ecología del Bentos Marino, por Manuel Vegas Vélez.
N° 10. Biosíntesis de Proteínas y el Código Genético, por Jorge E. Allende.
N° 11. Fundamentos de Inmunología e Inmunoquímica, por Félix Córdoba Alva y Sergio Estrada Parra.
N° 12. Bacteriófagos, por Romilio Espejo T.
N° 13. Biogeografía de América Latina, por Angel L. Cabrera y Abraham Willink.
N° 14. Relación Hospedante-Parásito. Mecanismo de Patogenicidad de los Microorganismos, por Manuel Rodríguez Leiva.
N° 15. Genética de Poblaciones Humanas, por Francisco Rothhammer.
N° 16. Introducción a la Ecofisiología Vegetal, por Ernesto Medina.
N° 17. Aspectos de Biología Celular y la Transformación Maligna, por Manuel Rieber.
N° 18. Transporte a Través de la Membrana Celular, por P. J. Garrahan y A. F. Rega.
N° 19. Duplicación Cromosómica y Heterocromatina a Nivel Molecular y Citológico, por Nestor O. Bianchi.
N° 20. Citogenética Básica y Biología de los Cromosomas, por Francisco A. Sáez y Horacio Cardoso.
N° 21. Ecología de Poblaciones Animales, por Jorge E. Rabinovich.
N° 22. Metodología para el Estudio de la Vegetación, por Silvia D. Matteucci y Aída Colma.
N° 23. Los Sistemas Ecológicos y la Humanidad, por Ariel E. Lugo y Gregory L. Morris.
N° 24. A Germinação das Sementes, por Luiz Gouvêa Laboriau.
N° 25. Introducción a la Farmacocinética, por Edison Cid Cárcamo.
N° 26. Introducción a la Teoría y Práctica de la Taxonomía Numérica, por Jorge Víctor Crisci y María Fernanda López Armengol.
N° 27. ¿Qué es la Diferenciación Celular?, por Roberto B. García y Susana Pereyra Alfonso.

N° 28. Limnología Sanitaria. Estudio de la Polución de Aguas Continentales, por Samuel Murgel Branco.
N° 29. Etología: El Estudio Biológico del Comportamiento Animal, por Raúl Vaz-Ferreira.
N° 30. Fotosíntesis, por Carlos S. Andreo y Rubén H. Vallejos.
N° 31. Pesca y Piscicultura en Aguas Continentales de América Latina, por Argentino A. Bonetto y Hugo P. Castello.
N° 32. Fundamentos de Genética Biométrica y sus Aplicaciones al Mejoramiento Genético, por Jorge A. Mariotti.

En preparación

Serie de matemática

Geometrías Finitas, por Oscar Barriga
Computadoras y Procesamiento de Datos, por Julio Villanueva y Oscar Harasic.
Principios Matemáticos da Dinâmica dos Fluidos, por Guilherme M. de la Penha.
Análisis Multivariado-Método de Componentes Principales, por Laura Pla.

Serie de física

Teoría de Fluidos en Equilibrio, por Antonio E. Rodríguez y Roberto E. Caligaris.

Serie de biología

Fitomorfología Funcional y Adaptativa, por Elías R. de la Sota.
Origen y Anatomía del Cromosoma Eucarionte, por Nestor O. Bianchi.
Limnología Básica, por José Galizia Tundisi.
El Plancton de las Aguas Continentales, por Aída González de Infante.

Nota: Las personas interesadas en adquirir estas monografías deben dirigirse a la Oficina de Ventas y Promoción, Departamento de Información Pública, Organización de los Estados Americanos, Washington, D.C., 20006-4499 o a las Oficinas de la OEA en el país respectivo.